(continued from part 23)

Looking at *figure 6*, we can see that current I_{in}, flows through resistor R_1 to junction P, and current I_f flows through resistor R_f. The ampifier's impedance is high enough to make the current I_- that flows into the inverting input negligible, compared with currents I_{in} and I_f, so that $I_{in} = I_f$. If we also assume that the current I_+ into the non-inverting input is negligible, then the non-inverting input is at 0 V, and the voltage at the inverting input is V, with respect to ground. This gives us:

$$I_{in} = \frac{V_{in} - V}{R_1}$$

and:

$$I_f = \frac{V - V_{out}}{R_f}$$

Remember, $I_{in} = I_f$ so:

$$\frac{V_{in} - V}{R_1} = \frac{V - V_{out}}{R_f}$$

If we substitute $-V_{out}/A_{OL}$ for V, since V is the inverting input voltage and the non-inverting input voltage is zero, then:

$$\frac{V_{in} + V_{out}/A_{OL}}{R_1} = \frac{-V_{out}/A_{OL} - V_{out}}{R_f}$$

rearranging this equation gives us:

$$\frac{V_{in}}{R_1} = \frac{-V_{out}}{R_1 A_{OL}} - \frac{V_{out}}{R_f A_{OL}} - \frac{V_{out}}{R_f}$$

so:

$$\frac{R_f}{R_1} V_{in} = -V_{out}\left(\frac{R_f}{R_1 A_{OL}} + \frac{1}{A_{OL}} + 1\right)$$

The closed loop gain, G, can therefore be seen to be:

$$G = \frac{V_{out}}{V_{in}}$$

$$= \frac{-R_f/R_1}{1/A_{OL}\,(R_f/R_1 + 1) + 1}$$

Now, if A_{OL} is large enough, we can ignore the term multiplied by $1/A_{OL}$ (since this will be a very small number) giving a closed loop gain of:

$$G = \frac{V_{out}}{V_{in}} = \frac{-R_f}{R_1}$$

(This gives us a negative value, because the inverting input is being used.) This proves, therefore, that the gain, G, is determined by the ratio of the resistors R_f and R_1.

There is an alternative analysis to the rather long, but straightforward proof that has just been explained. We'll take a look at it now, as it introduces an important point about this type of circuit.

If the amplifier gain A_{OL} is very large, then, since $V = V_{out}/A_{OL}$, V must be very small compared with V_{out}. Thus, the inverting input can be thought of as being at 0 V, if A_{OL} is large enough. Therefore, the point P in *figure 6*, whilst not directly connected to ground, can be thought of as being virtually at 0 V and is therefore termed a **virtual earth**.

As a practical example, we'll assign values to the resistors in *figure 6*. We want the circuit to give a gain of –10 (a negative, remember, because the output will be inverted), and we define R_1 to be 100 kΩ.

$$-G = \frac{R_f}{R_1}$$

so $R_f = 10 \times 100\ \text{k}\Omega = 1\ \text{M}\Omega$. Resistor R_2 is defined by:

$$R_2 = \frac{R_f R_1}{R_f + R_1}$$

$$= \frac{1 \times 10^6 \times 1 \times 10^5}{1 \times 10^6 + 1 \times 10^5}$$

$$= \frac{1 \times 10^{11}}{11 \times 10^5}$$

$$= 90909.\dot{0}\dot{9}\ \Omega$$

The nearest preferred value (real value) of a resistor to this is 100 kΩ, so:

$$R_2 = 100\ \text{k}\Omega$$

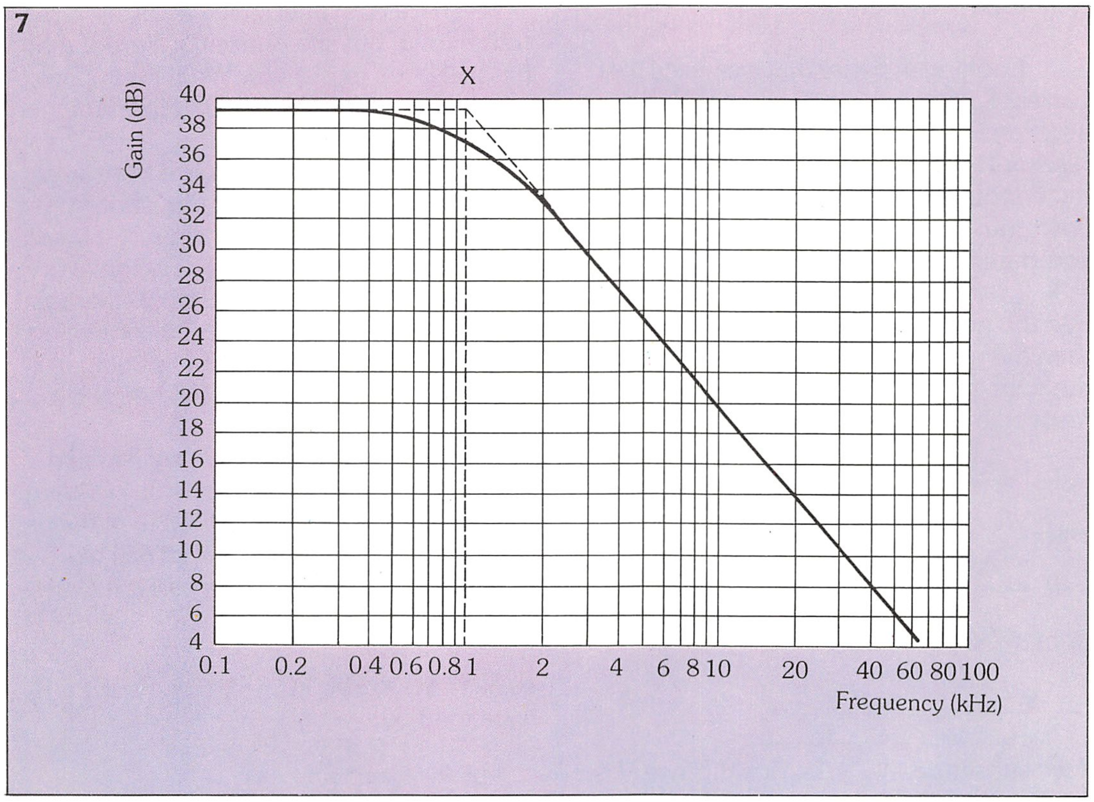
We can therefore see that if a voltage of 10 mV is applied to the input of this circuit, the output will be:

$$10 \times -10\ \text{mV} = -100\ \text{mV}$$
$$= -0.1\ \text{V}$$

Gain and frequency

Now we have seen how to control an op-amp's gain using negative feedback, let's return to the earlier problem of frequency response and bandwidth. We know that the amplifier's gain drops rapidly with frequency (*figure 4b*) and we have assumed the gain in our examples to be due to low frequency input signals. To recap, we know that the bandwidth is defined as the area of the frequency response over which the gain does not fall by more than 3 dB.

The decibel is a measure that enables us to express the ratio of two powers: this is



7. Open loop gain *vs* frequency, showing corner frequency at point X.

8. Open and closed loop gains plotted against frequency showing the greater bandwidth of the closed loop gain.

9. (a) Open loop amplifier circuit; **(b)** response curve.

especially important when working with amplifiers as it is much more useful to know the comparative or relative gains, rather than the actual gains. This enables us to compare the qualities of different amplifier circuits, without having to work out gain figures for specific values, that would have to be identical in each case.

So, to express the ratio of two voltages (say input voltage and output voltage, in this case) we take the logarithm to the base 10 of the ratio, and multiply this logarithm by 20. The gain in decibels is therefore found by the formula:

$$\text{Gain (dB)} = 20 \log_{10} \frac{V_{out}}{V_{in}}$$

The graph in *figure 7* shows an amplifier's open loop gain in dB, plotted against frequency (on a logarithmic scale). As you can see, point X indicates the device's bandwidth, and is known as the **corner frequency**. *Figure 8* illustrates the way that the bandwidth of the device is increased by the reduction in gain. We know that the open-loop gain, A_{OL}, will give us a bandwidth equal to frequency f_2 in *figure 8*. A closed loop gain, G, on the other hand, gives a greater bandwidth – it has, in fact, been increased by the feedback factor $(1 + \beta A_{OL})$ and is shown by the lower curve in *figure 8*. The gain and the bandwidth are thus linked by the gain bandwidth product of the amplifier which is equal to:

$$\frac{A_{OL}}{1 + \beta A_{OL}} \times f_2 (1 + \beta A_{OL}) = A_{OL} f_2 = f_3$$

You'll notice that the gain bandwidth product is constant, and does not depend on the feedback factor. If the feedback factor is increased (by increasing β), then the gain, G, will decrease – but the bandwidth will increase by the same proportion. The general rule is that:

$A_{OL} \times f_2 = X$ Hz (point f_3 in *figure 8*)

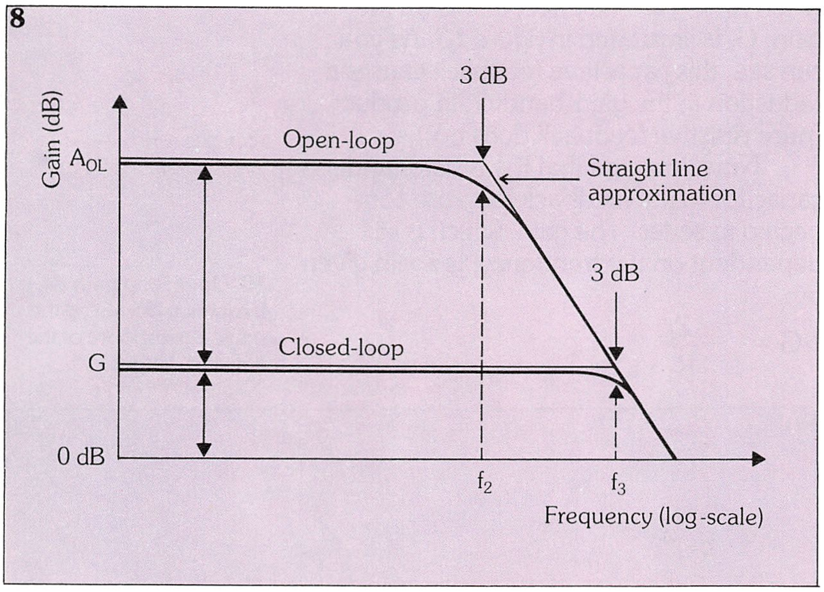
As an example, we'll assign some values to *figure 8*:

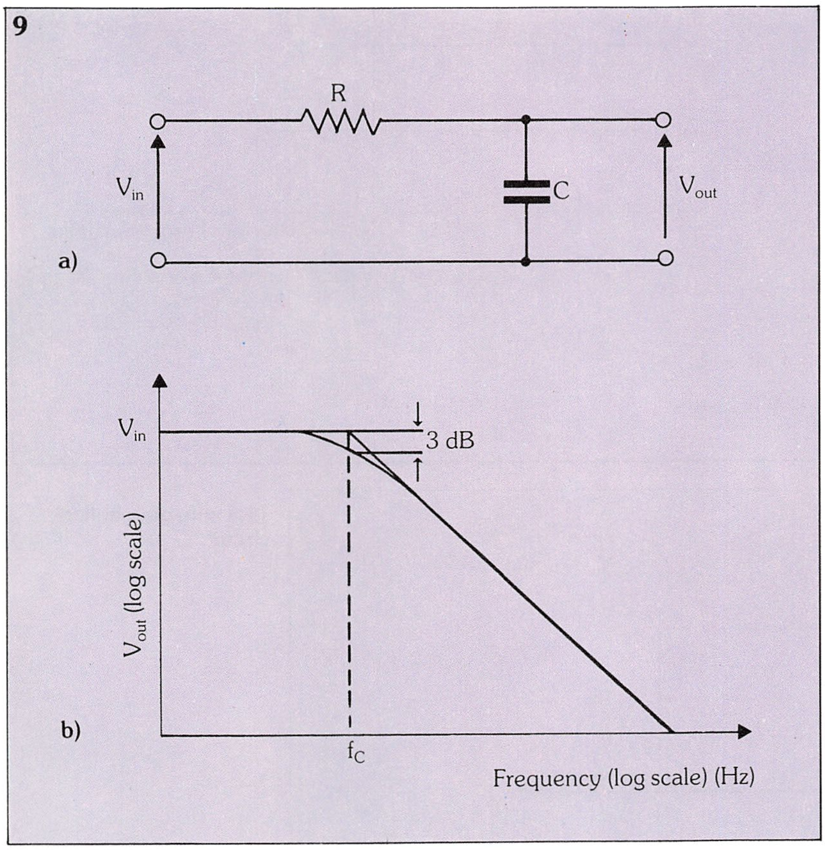
A_{OL} = 60 dB
f_2 = 1 kHz

To calculate the feedback factor we have to obtain a real value for the gain and this is equal to the antilog (10^x) of 60/20, which is antilog 3 or 10^3. The gain bandwidth product is 1 kHz $\times 10^3$ or 1×10^6.

If we want the new bandwidth to be 40 kHz, then the new gain, G, is equal to the gain bandwidth product divided by 40, therefore:

$$G = \frac{1 \times 10^6}{40 \times 10^3}$$





$\therefore G = 25$
and $\log_{10} 25 \times 20 = 28$ dB. So the gain, G, that gives us a bandwidth of 40 kHz is 28 dB.

We know that the open loop amplifier circuit is only really effective at low frequencies – it therefore acts like an amplifying low-pass filter which, as its name suggests, only conducts low frequency signals.

The circuit diagram is shown in *figure 9a.* As you can see from the response curve in *figure 9b,* the corner frequency, f_c, is equivalent to the bandwidth of the amplifier circuit. This corner frequency:

$$f_c = \frac{1}{2\pi RC}$$

So, at f_c, $V_{out} = 0.707 \times V_{in}$ where 0.707 in this case corresponds to -3 dB.

The drop in output or gain for this type of circuit corresponds to 20 dB for every decade increase in frequency above the corner frequency. That is to say, that when the frequency becomes ten times larger, the output/gain drops by 20 dB. This well known response is known as a **single lag** response, our op-amp then is a single lag amplifier. Where two single lag circuits are connected in series, the drop is 40 dB per decade which is known as a **double lag** response, and so on. *Figure 10* illustrates the single lag response of our amplifier's output.

Unity gain buffer

What would happen to the circuit in *figure 9a* if the negative feedback resistor was omitted, and the loop connected direct? Such a circuit (shown in *figure 11*) is known as a **unity gain buffer** or **voltage follower,** for reasons that will become apparent.

You'll notice that the non-inverting input is now being used as the V_{in} input. The overall system input is V_{in}, but as the op-amp is now being used as a differential amplifier, the overall input is $V_+ - V_-$; the output of the circuit is equal to this input multiplied by A_{OL}. As V_+ is equal to V_{in}, and V_- is equal to V_{out}, the input to the differential amplifier is the difference between the input and the output ($V_{in} - V_{out}$). The output is therefore A_{OL} times this value:

$$V_{out} = A_{OL}\,(V_{in} - V_{out})$$

which can be written as:

$$(1 + A_{OL})\,V_{out} = A_{OL}\,V_{in}$$

So:

$$V_{out} = \left(\frac{A_{OL}}{1 + A_{OL}}\right) V_{in}$$

As we have noted previously, if A_{OL} is sufficiently large (usually about 200,000) then the 1 can be ignored, thus the output

is equal to the input. This means that the circuit has a gain of 1, i.e. **unity gain.**

What, then, is the point of the circuit if it's output is the same as the input it received? One of the characteristics of op-amps is that they present a large input impedance and a low output impedance, thereby acting as an isolating network that effectively separates a high impedance source from a low impedance load. Such a device, of course, is known as a **buffer.**

Frequency dependent feedback

A virtual earth amplifier with a feedback impedance Z_f, comprising a capacitor C_f in parallel with a resistor R_f, is shown in *figure 12.* We know that the gain of this negative feedback circuit is given by:

$$G = -\frac{Z_f}{R_l}$$

The magnitude of the impedance of a capacitor and resistor in parallel can be found from phasor addition:

$$\frac{1}{Z_f} = \sqrt{\frac{1}{R_f^2} + \frac{1}{X_C^2}}$$

where X_C is the capacitance reactance of C_f. This is, of course, dependent on frequency.

The frequency response curve of this circuit is shown in *figure 13.* The frequency of the corner point and hence the bandwidth is defined by:

$$f_{corner} = \frac{1}{2\pi R_f C_f}$$

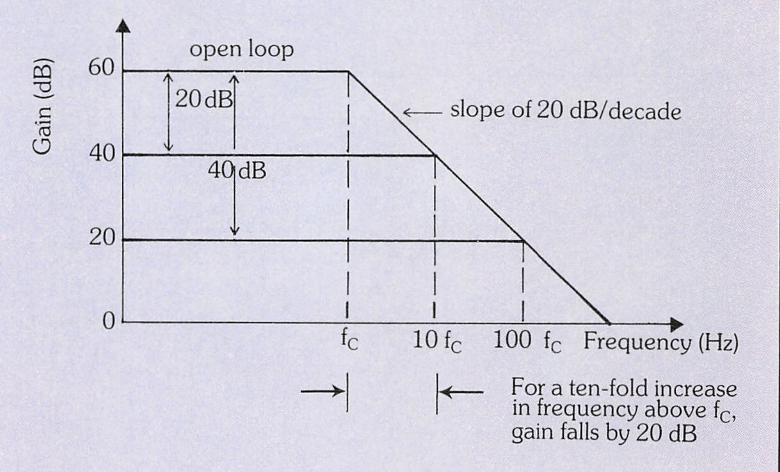
when $X_C = R_f$. The effect of this on the gain, G, is illustrated in *figure 13.* As you can see, this capacitive feedback causes a reduction in the gain-bandwidth product (pure resistive feedback does not).

Now, imagine that the resistor and capacitor in the feedback loop are connected in series. The gain, which is still dependent on the frequency, is again given by:

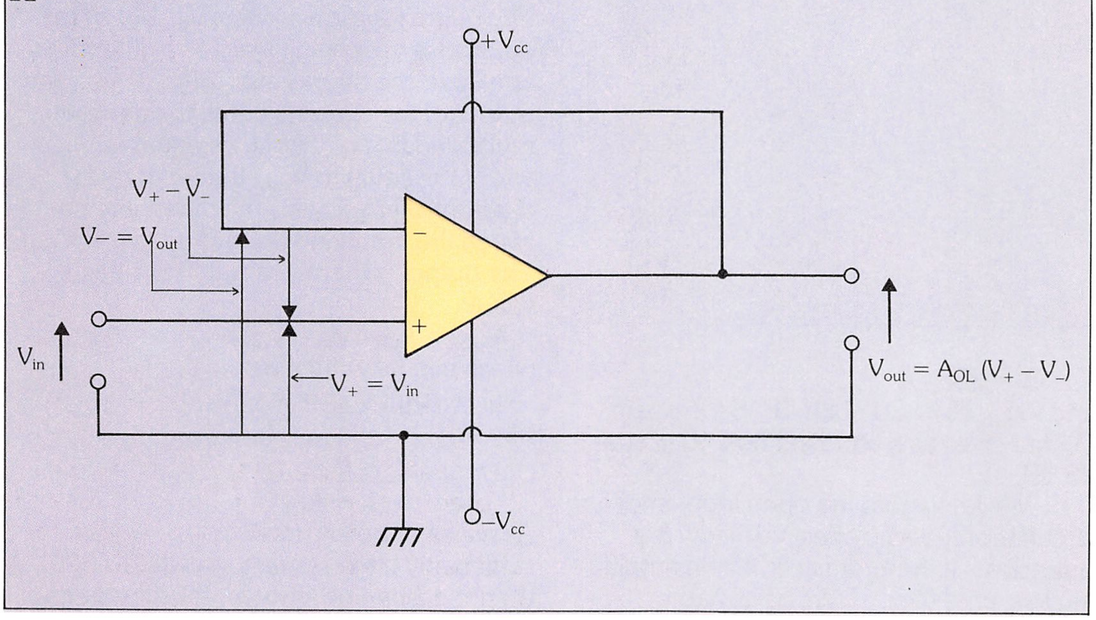
$$G = -\frac{Z_f}{R_1}$$

10. Open loop gain *vs* frequency illustrating the single lag response of the amplifier output.

10



11



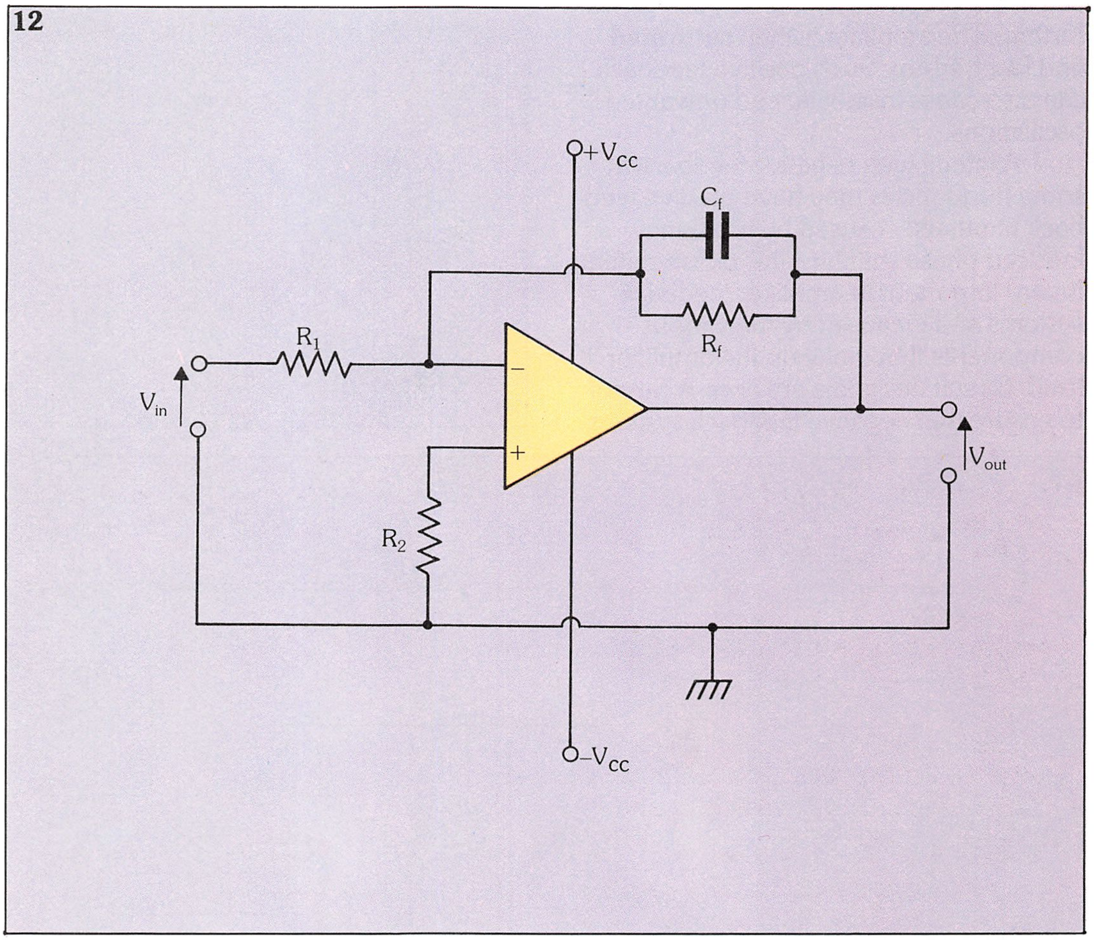
11. Unity gain buffer circuit.

The magnitude of the impedance of the capacitor and resistor in series is found by phasor addition:

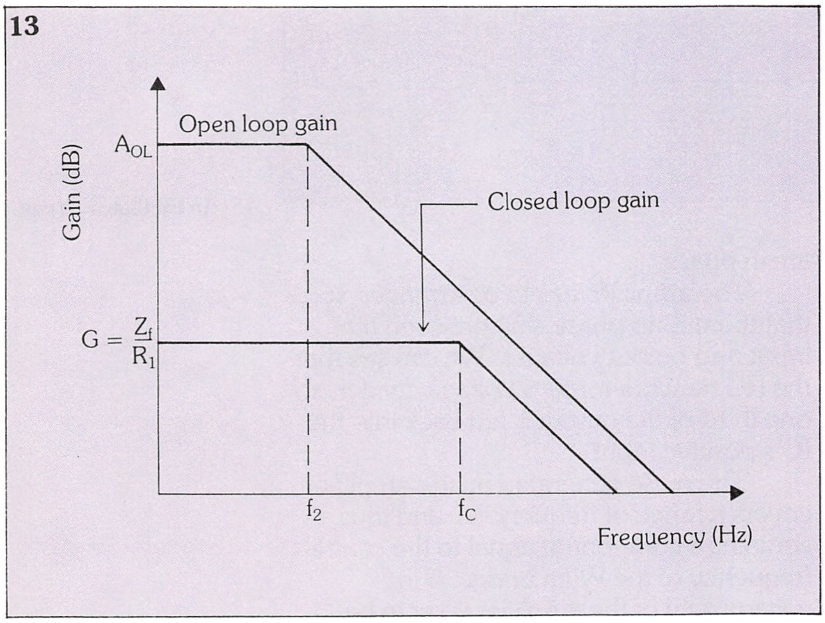
$$Z_f = \sqrt{R_f^2 + X_C^2}$$

At high frequencies, the value of X_C is negligible, leaving R_f to make up the impedance Z_f. But at low frequencies, the value of X_C rises, thus decreasing the amount of negative feedback that will pass

12. A negative feedback virtual earth amplifier circuit.



13. Frequency response curve for the circuit of *figure 12*.



along the loop. The frequency response curve therefore looks like that in *figure 14*. As you can see, a large low frequency boost flattens out in the mid-range and tails off normally into high frequencies.

The corner frequency, f′, is found in the normal way, using the formula:

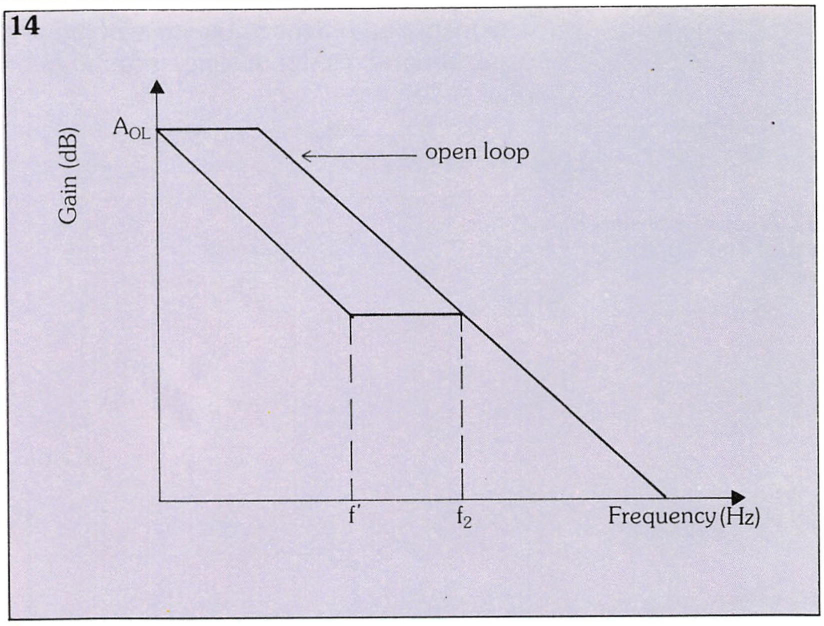
$$f' = \frac{1}{2\pi R_f C_f}$$

In real amplifiers, the low frequency gain and the corner frequency are often ill-defined. Using capacitive feedback, on the other hand, these factors can be well defined as they only depend on R_f, C_f and R_1. This ensures that low frequency gain and bandwidth are firmly under the control of the designer.

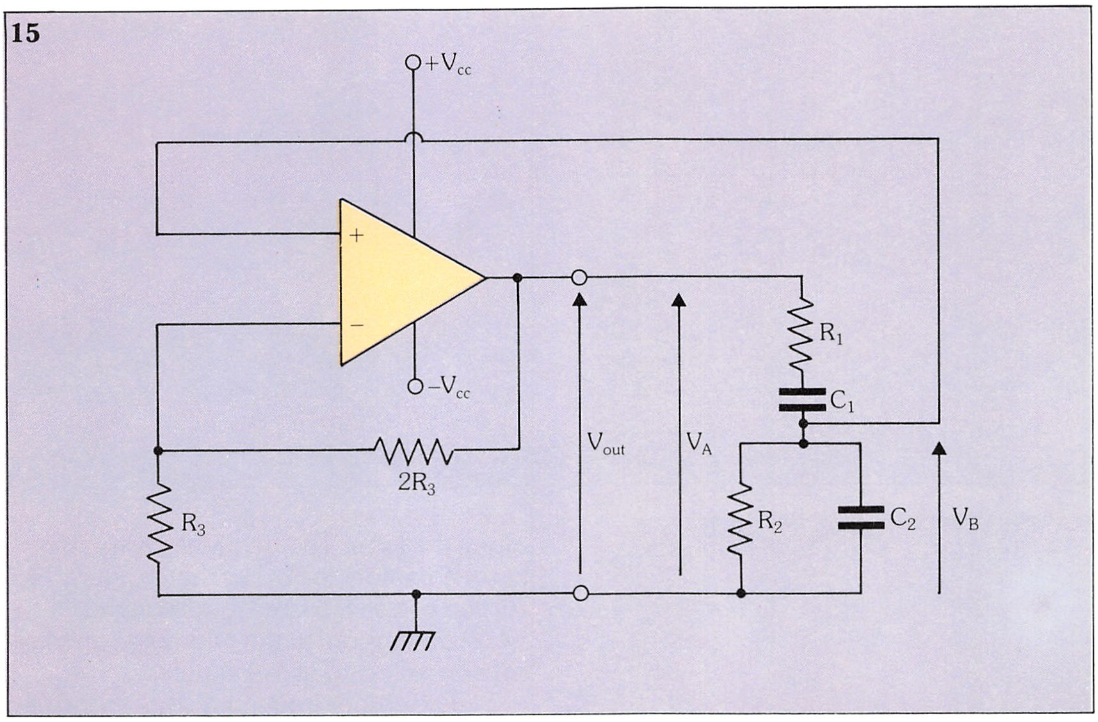
Positive feedback

If the output of an op-amp is directly fed back into its non-inverting input, the electrical noise generated by the amplifier will cause the output to rise, until it can go no further. The amplifier is then **saturated** and **latched-up**. Such positive feedback can also cause instability and unwanted oscillations.

A circuit with negative feedback at some frequencies may have positive feedback at others – caused by frequency induced phase shifting. This phase shifting is very important in amplifier feedback systems and is caused by the various components that make up the amplifier itself. Circuit designers are always careful to ensure that negative feedback systems do not turn positive at some frequency, thus causing oscillations and instability resulting in circuit malfunction.



14. Response curve for a negative feedback circuit where the capacitor and resistor are connected in series.



15. An oscillator circuit.

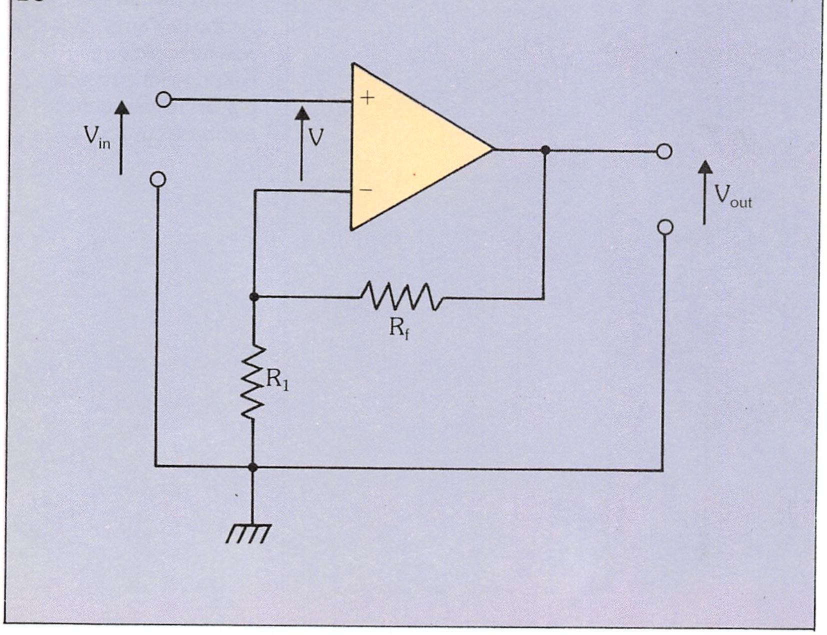
Op-amp oscillator

Figure 15 illustrates a popular oscillator circuit. You'll notice that there is an RC network to the right of the IC – this is connected in an arrangement known as a **Wien bridge.** This circuit will only oscillate at one specific centre frequency; the frequency at which the voltages V_A and V_B are in phase.

The amplifier has to be arranged so that there is no phase shift between the input and output voltages. We can see that the RC network forms a voltage divider, so one third of the output is fed back into the IC's positive input.

The noise generated by the amplifier covers a range of frequencies, and thus contains a component equal to the centre frequency of the Wien bridge. If the voltage gain of the amplifier is set to be

16



16. A basic non-inverting amplifier circuit.

slightly more than 3, then the component at the centre frequency will start to go round the circuit, becoming larger each time it completes the loop. This continues until the amplifier reaches saturation. However, an amplifier's gain is reduced when it reaches saturation because the output voltage increases less than it would if the amplifier was not saturated. So the gain, which was just higher than 3, drops back slightly to 3 ensuring that the output sine wave stays constant in amplitude.

Non-inverting amplifier

One application of op-amps that demands negative feedback and (usually) a non-inverted output is the audio amplifier: the input signal is applied to the non-inverting input, while a proportion of the output is fed back to the inverting input. This has the effect of increasing bandwidth and keeping the output in phase.

In audio work, it is obviously important to try and keep the frequency response characteristics as close as possible to the range of human hearing. (A healthy young ear should be capable of detecting sounds in the frequency range 20 Hz to 20 kHz.)

Figure 16 shows a basic non-inverting amplifier circuit. The output:

$$V_{out} = A_{OL}\,V$$

while:

$$V = V_{in} - \left(\frac{R_1}{R_1 + R_f}\right)V_{out}$$

R_1 and R_f act as a voltage divider across the output. Since:

$$V_{out} = A_{OL}\,V$$

$$= A_{OL}\,V_{in} - \left(\frac{A_{OL}\,R_1\,V_{out}}{R_1 + R_f}\right)$$

then:

$$V_{out} = \frac{A_{OL}\,V_{in}}{1 + A_{OL}\,[R_1 / (R_1 + R_f)]}$$

$$= \frac{V_{in}}{1/A_{OL} + R_1/(R_1 + R_f)}$$

The non-inverting amplifier circuit's gain is therefore:

$$G = \frac{V_{out}}{V_{in}}$$

$$= \frac{1}{1/A_{OL} + R_1/(R_1 + R_f)}$$

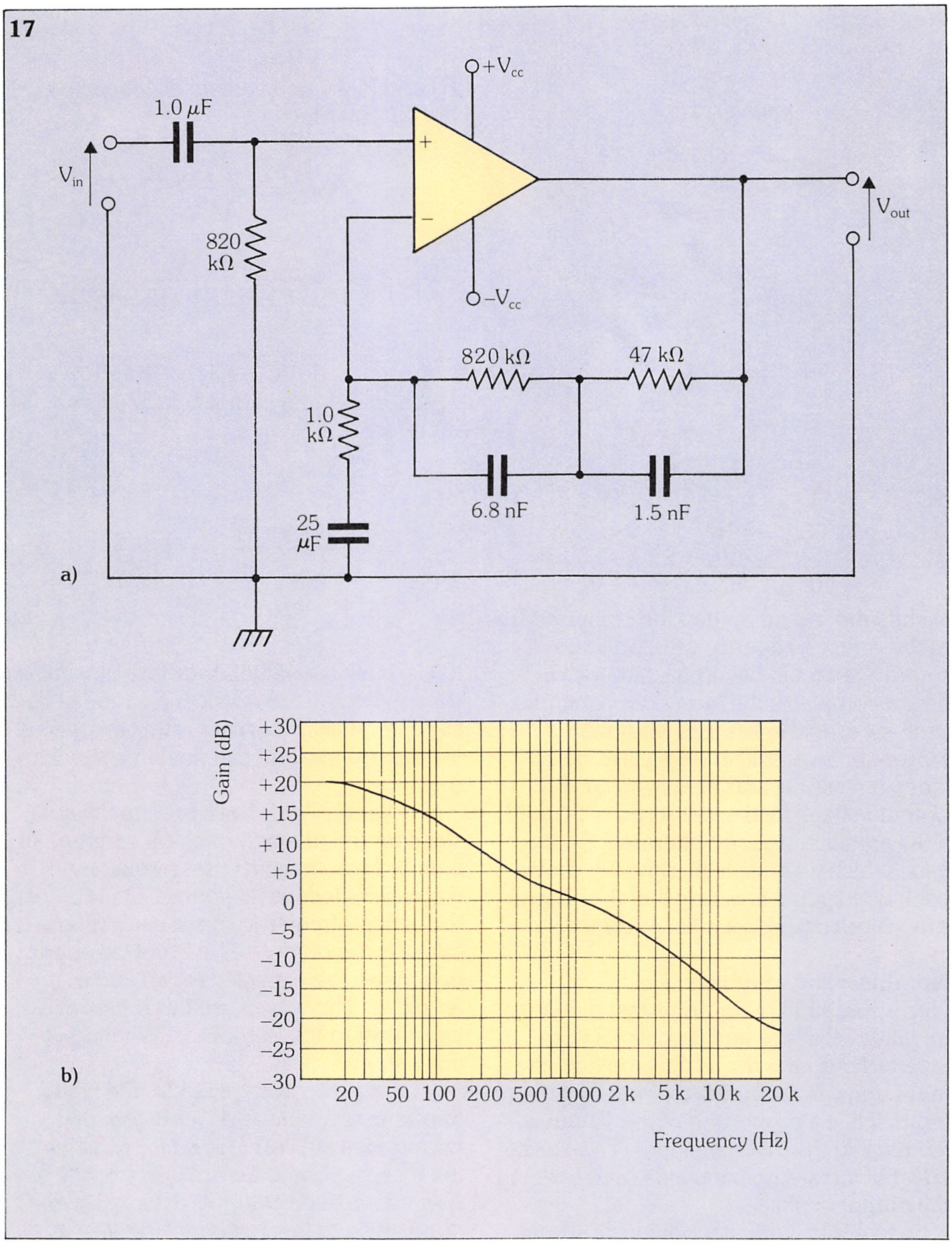
$$\simeq \frac{R_1 + R_f}{R_1}$$

if A_{OL} is a large value. As before, the gain is determined by the resistors R_1 and R_f. The big difference in this case is that the gain is always positive and can never be less than one.

Figure 17a shows a pre-amplifier stage for the magnetic pick-up cartridge of a record player which uses frequency dependent feedback; *figure 17b* indicates the circuit's frequency response. As you can see, the output curve is not the shape you may have expected for an audio amplifier. This is because this pre-amp is **equalised** to IEC standards. What does this mean?

When records are made, the high frequencies are boosted while the low frequencies are cut. This is to ensure that the record grooves do not become deformed, through attempts at carrying excessively loud low frequency signals. In order to ensure that the amount of bass cut and treble boost are consistent in recordings world-wide, an international standard was set up by the International Electrotechnical Commission (IEC). It stands to reason that whenever we replay a record we must counteract or equalise this tonal imbalance. The same IEC standard thus applies to the characteristics of all magnetic pick-up pre-amplifiers world-wide.

The RC networks in the circuit's



17. (a) Circuit diagram for the pre-amp stage of a magnetic pick-up cartridge for a record player; **(b)** frequency response curve.

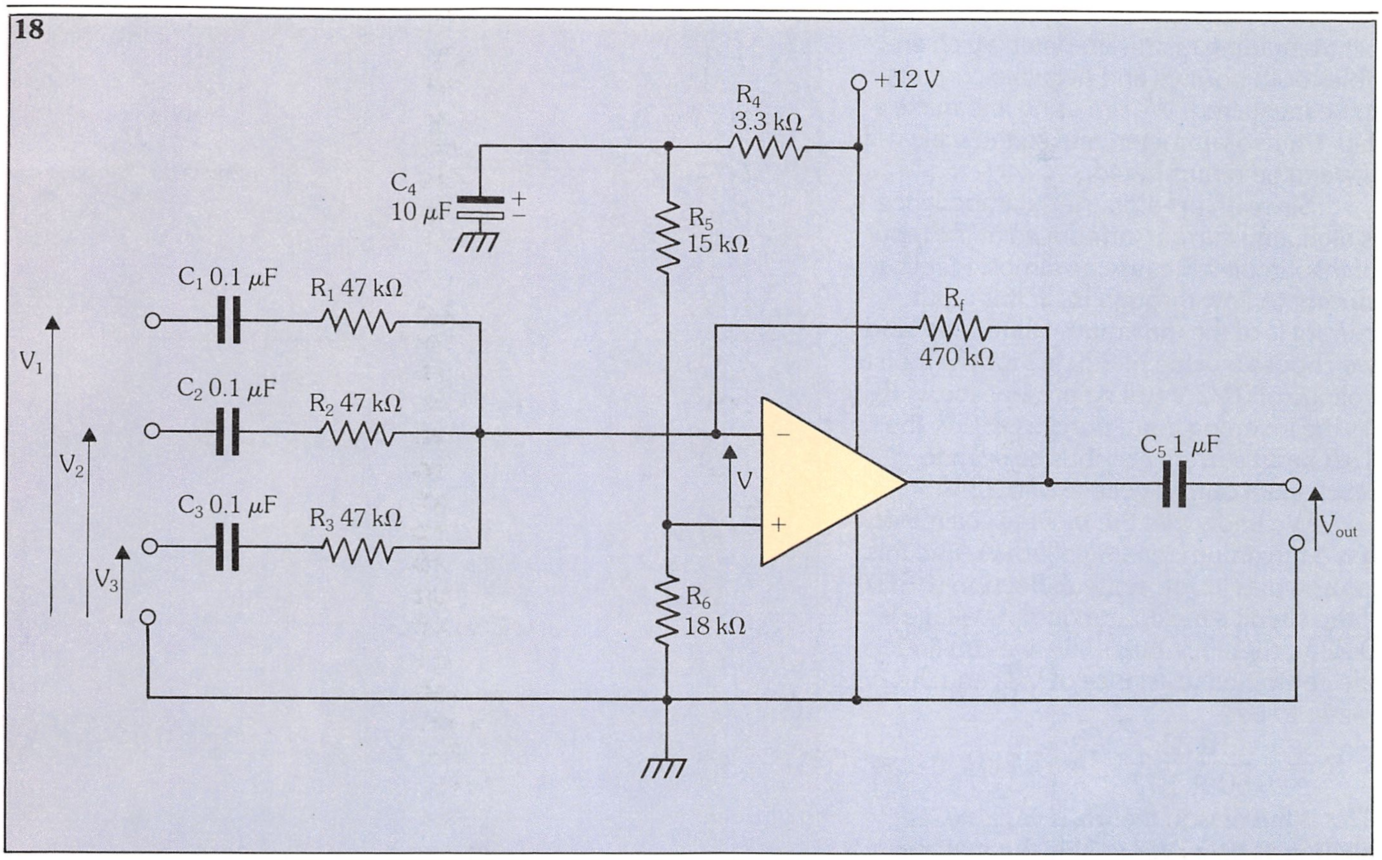
negative feedback loop (*figure 17a*) provide the bass boost and treble cut required to give an output signal for power amplification.

With a little imagination, you can see how the frequency response curves in *figures 13* and *14* would give an IEC response when combined. Remember, the straight line curves that we have been using have been approximations and not entirely accurate. Actual amplifier response characteristics are, of course, more suited to the production of the IEC curve – which is why it is the shape it is!

Summing amplifier

The **summing amplifier** is a classic arithmetic op-amp circuit, but is also used in non-computational applications. *Figure 18* illustrates a summing amplifier used as an

18



18. Summing amplifier circuit used as an audio mixer.

audio mixer. This gives an output signal that is an amplified combination of its different inputs. In this case, the output:

$$V_{out} = -\left[V_1\left(\frac{R_f}{R_1}\right) + V_2\left(\frac{R_f}{R_2}\right) + V_3\left(\frac{R_f}{R_3}\right)\right]$$

So we can see that the output of this circuit is the weighted sum of its inputs. The general equation for circuits of this type is:

$$V_{out} = -\left[\left(\frac{R_f}{R_1}\right)V_1 + \left(\frac{R_f}{R_2}\right)V_2 + \ldots + \left(\frac{R_f}{R_n}\right)V_n\right]$$

Looking at our circuit, we can see that the ratio R_f/R_n will give us a gain of 10. This value can, of course, be changed by altering the value of R_f. The input resistors are of the values chosen, to prevent any one input signal from interfering with the others.

Although this is an inverting circuit, its suitability for audio use is not unduly compromised by the phase shift. However, if this *is* a problem, then another inverting op-amp arranged to give unity gain can be attached to the circuit which restores the phase to its original condition.

Coupling capacitors

Most amplifying circuits (and many other circuits) have capacitors connected in series with their inputs and outputs. These are known as **coupling capacitors** and exist to prevent any DC signals within the circuit interfering with DC signals in any other circuits it may be connected to. Remember, capacitors appear to conduct AC while blocking DC.

This, however, has an effect on an amplifier's frequency response. The capacitor and op-amp (*figure 19a*) act as a **high-pass filter** with an equivalent circuit as in *figure 19b*, and giving a frequency response like that shown in *figure 19c*. This is known as a **low frequency roll-off.**

Op-amps in meters

Op-amps can be employed to make accurate and sensitive meters, useful for circuit testing and experimental work. *Figure 20* shows a 741 op-amp at the heart of such a circuit: a 1μA current meter. The basic principle of this circuit is that it will amplify a small input current sufficient to deflect a moving coil current meter. When calibrated, this provides an accurate metering system with a sensitivity unobtainable from a moving coil meter alone.

The actual meter used is a 50-0-50

centre reading micro-ammeter which enables both positive and negative currents to be measured. We are using it to make a 1-0-1 micro-ammeter, and so the scale will have to be renumbered.

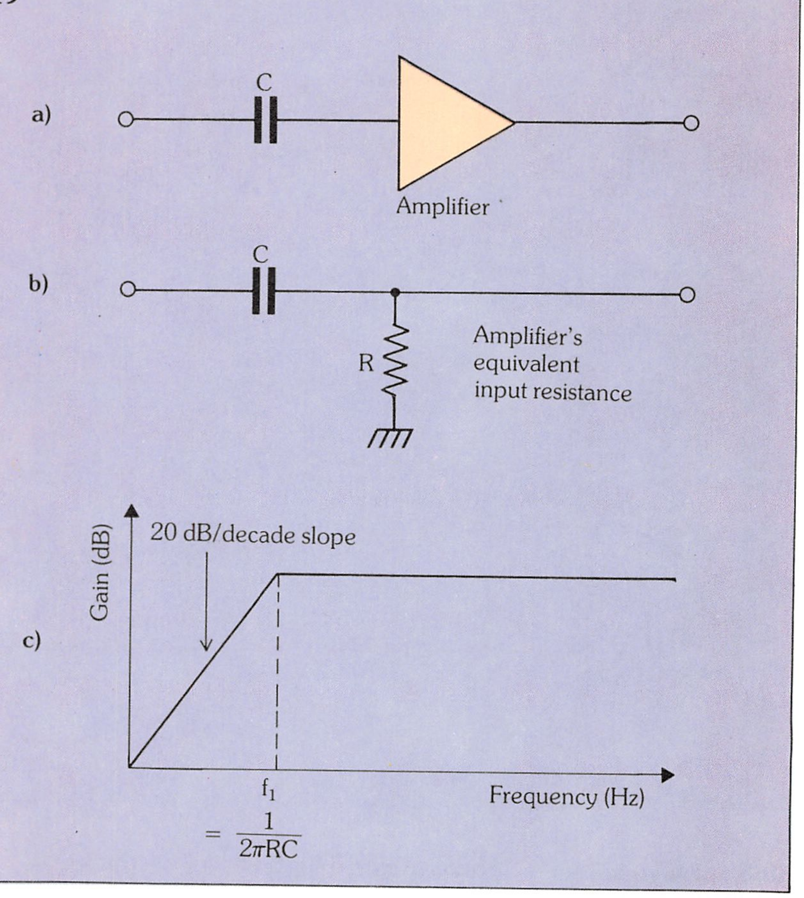
Since an op-amp's input impedance is high, any current introduced at the input of this circuit will cause an almost identical current to flow through R_3. If this input current is of the maximum value $1\mu A$, and we choose a value of 220 kΩ for R_3, then a voltage of 0.22 V will be present across R_3. As the inverting input remains at 0 V, the $1\mu A$ input current can thus be seen to result in an output voltage of 0.22 V.

We know that the moving coil meter has a maximum reading of $50\mu A$, and this is known as its **full scale deflection (FSD)**. If the circuit's maximum output voltage is 0.22 V, then by Ohm's law, we can find the combined resistance of VR_2 and the meter to be:

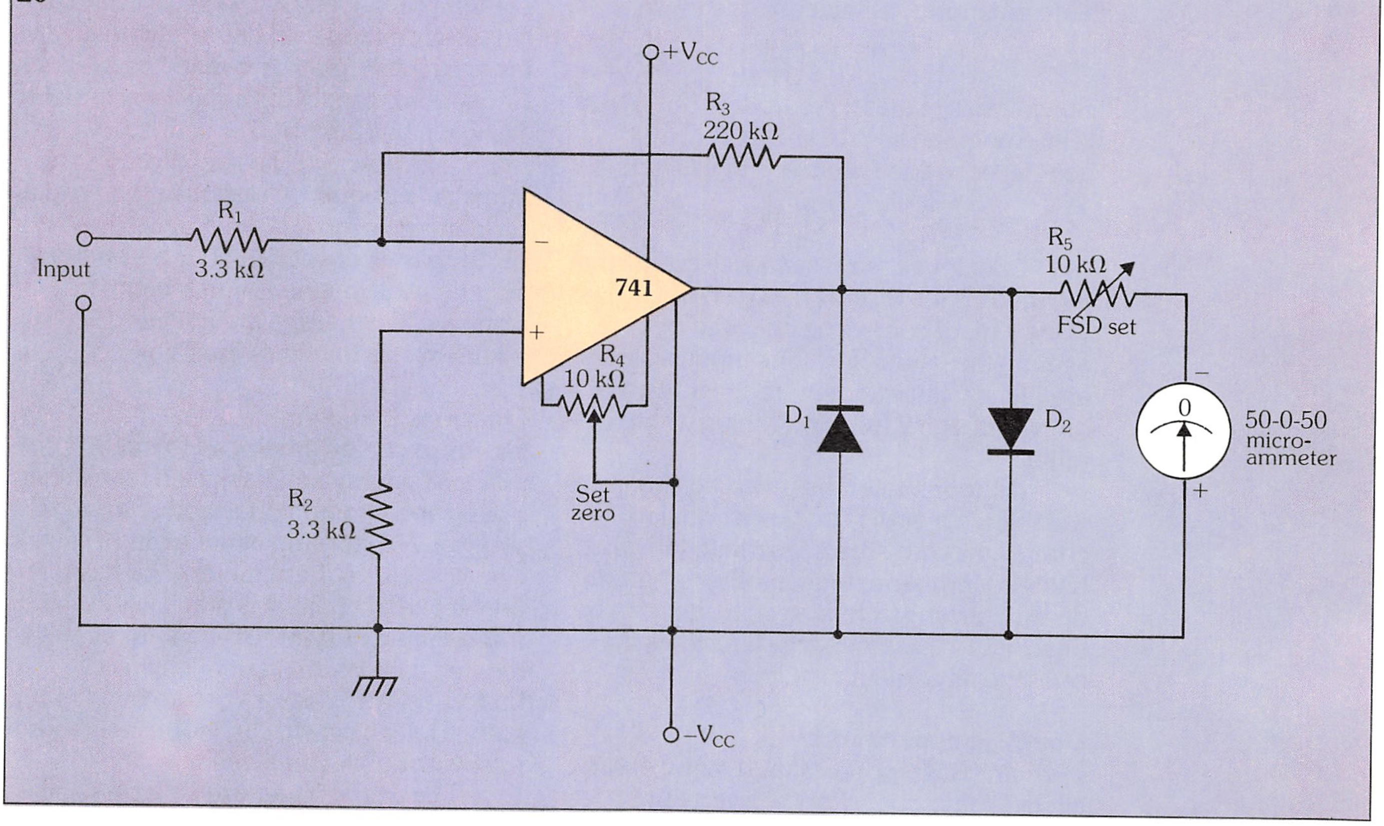
$$R = \frac{0.22}{50 \times 10^{-6}} = 4.4\ \text{k}\Omega$$

This is the reason that R_5 is variable, as altering its value will enable the meter to read maximum value (FSD) when the circuit has a $1\mu A$ input. This circuit can be

19



20



19. (a) A high-pass filter; (b) an equivalent circuit; **(c)** frequency response curve showing low frequency roll-off.

20. 741 op-amp used in a meter circuit.

redesigned (by altering component values) to measure different current ranges.

The two diodes in the circuit protect the meter from overloading, should the input current be too large. They will not conduct any noticeable current until the output voltage grows larger than 0.6 V.

You may be wondering about the function of the variable resistor marked **set zero.** Arguably, with no input, the output should also be zero. However, an op-amp's output voltage can drift with temperature, so, as the device warms up, the output moves away from zero thus affecting the meter reading. Provision is therefore made for a variable resistor to be included in op-amp circuits to ensure that the output is zero, when both inputs are zero. This is known as **offset null adjustment** or **offset nulling.**

Glossary

bandwidth — the frequency range of an amplifier over which the gain does not fall by more than 3 dB

closed loop amplifier — an amplifier circuit that employs a form of feedback in a loop from the output to the input

differential amplifier — op-amps can be used as differential amplifiers. These amplify the difference between the voltages present at the inverting and non-inverting input terminals

feedback factor — $1 + \beta A_{OL}$. When the gain of an amplifier without feedback is multiplied by the feedback factor, the gain with feedback is found. β is the fraction of the output voltage that is actually fed back to the input

gain — measure of the amount of amplification delivered by an amplifier. Is equal to the output divided by the input (measured either in voltage or current)

gain-bandwidth product — an amplifier's gain and bandwidth are linked by their product which, for a single lag amplifier, is constant and does not depend on the feedback factor.

inverting input — op-amp input marked '—' on circuit diagrams. Signal applied to this input will appear amplified and inverted at the output

latch-up — situation that occurs when an amplifier reaches saturation and the output reaches a constant maximum high value

non-inverting input — op-amp that enables an in-phase, amplified copy of the applied signal to be obtained as output

open loop amplifier — amplifier circuit with no feedback loop. The loop is thus open circuit, and the amplifier's gain and bandwidth are its 'natural' values

single lag — an op-amp's output will drop by 20 dB per decade increase in frequency. This is known as single lag. Two single lag circuits connected in series have a double lag and so on

virtual earth — point in a circuit that while not actually connected to earth, can be regarded as being at 0 V

Wien bridge — filter network made in bridge form from two resistors and two capacitors. This allows signals of only specific frequency to pass

ELECTRICAL TECHNOLOGY

Real and reactive power

In a circuit supplied by a steady (DC) voltage, the power consumed by a resistor is the product of voltage and current, i.e:

$$P = VI$$

If alternating current flows through the same circuit, then the power, p, at any instant of time is given by the same expression:

$$p = vi$$

where v and i are the instantaneous values of voltage and current in the circuit.

Power in a pure resistance

We can draw the voltage and current waveforms for the resistance in the circuit shown in *figure 1a,* which is supplied by a sinusoidal voltage $v = \hat{V} \sin \omega t$: these waveforms are shown in *figure 1b,* along with the instantaneous power, p. We can see that p is always positive, fluctuating between its maximum value and zero, twice in each voltage cycle. The maximum value of the power is equal to $\hat{V}\hat{I}$, and the average value is given by:

$$\begin{aligned} P &= \tfrac{1}{2}\hat{V}\hat{I} \\ &= \hat{V}\hat{I} \end{aligned}$$

where V and I are the rms values of the voltage and current.

Since V = RI, the above expression becomes:

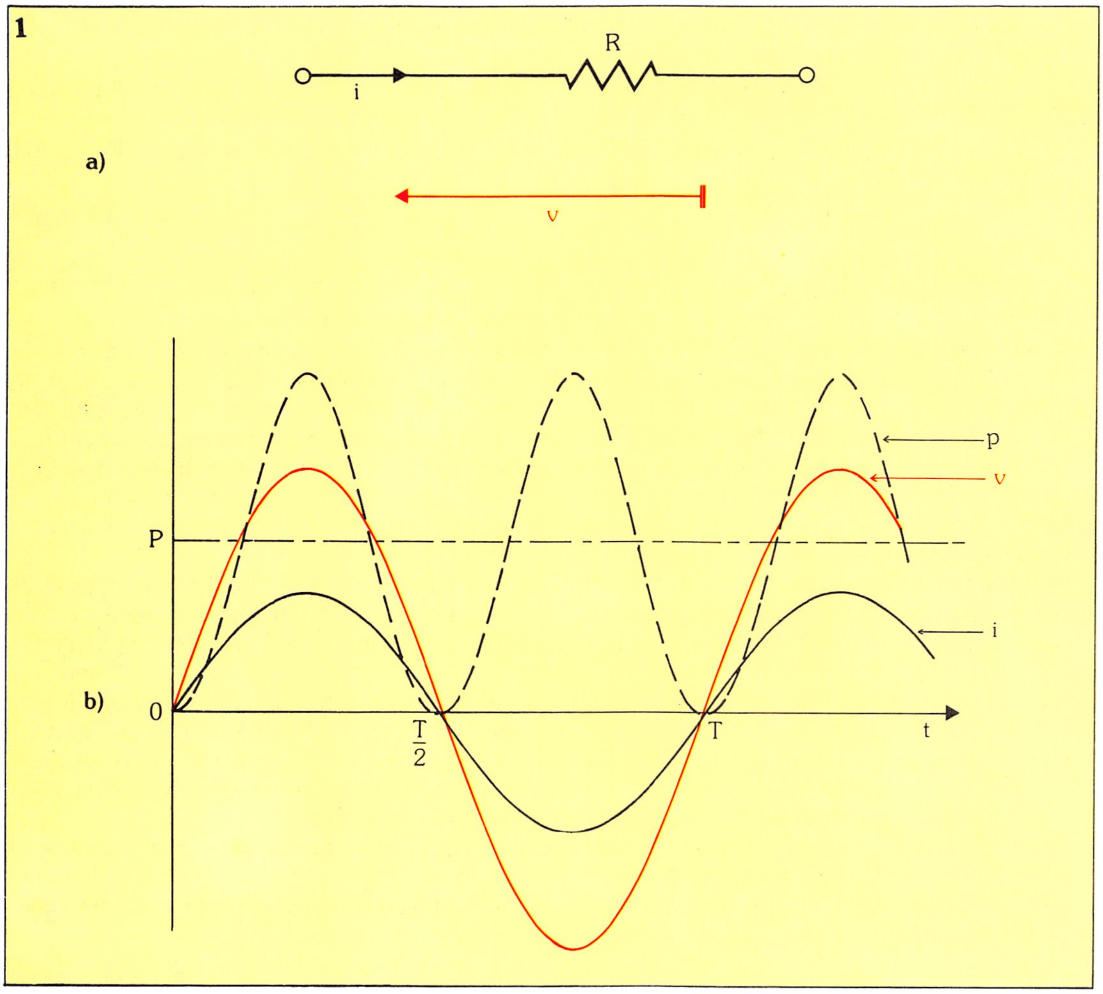
$$P = I^2R$$

Power in a pure capacitor

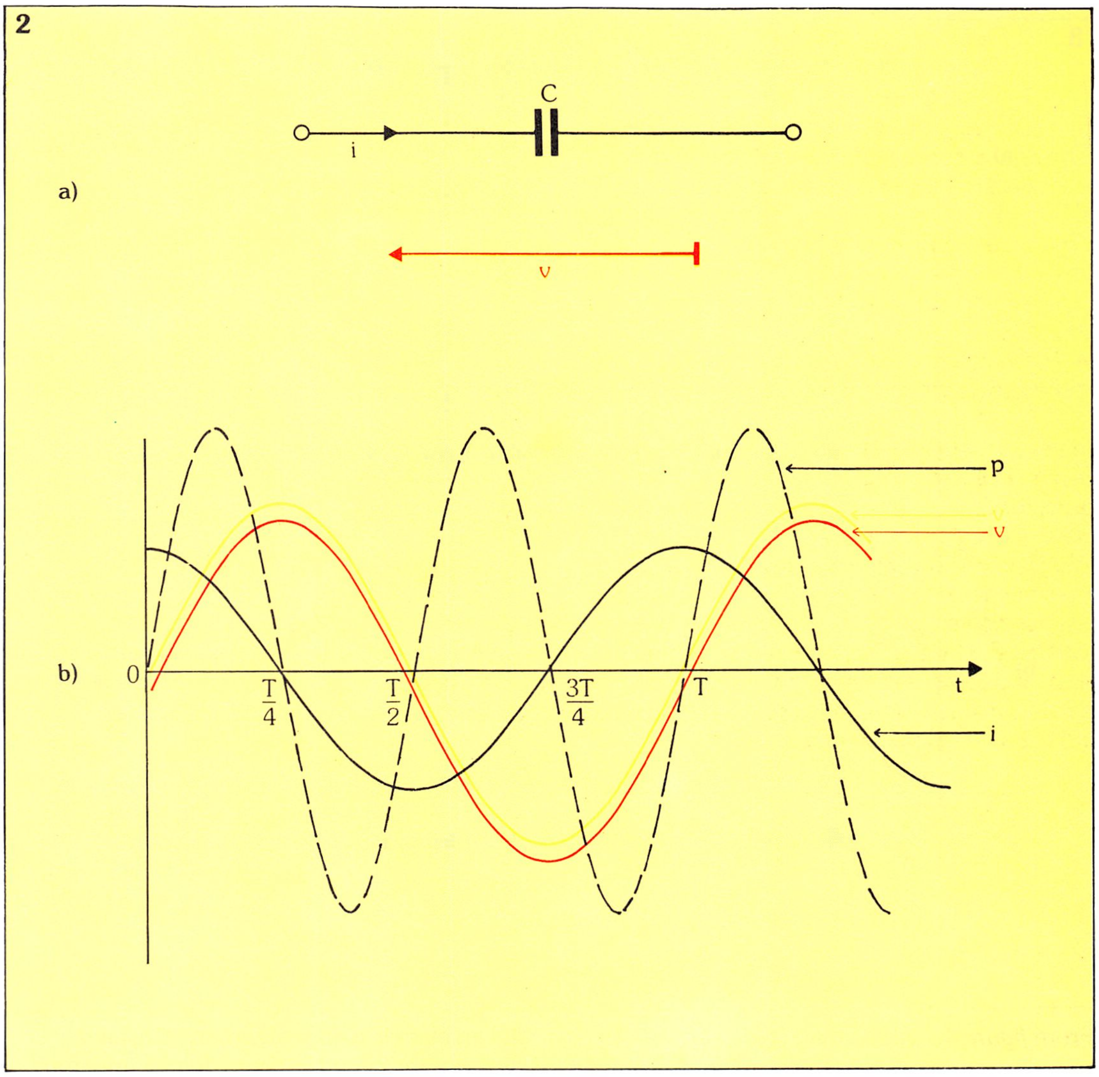
Now we'll consider the instantaneous power in a pure capacitor as shown in *figure 2a.* The current and voltage waveforms are shown in *figure 2b,* and the instantaneous power is obtained from their product:

$$p = vi$$

Once again, we see that power fluctuates at twice the frequency of the supply voltage. The average power over a long period of time, however, is zero, since from time zero to T/4 energy enters the capacitor (where the instantaneous power is positive), then from time T/4 to T/2 the same amount of energy leaves the



1. (a) Circuit supplied by a steady DC voltage; **(b)** voltage and current waveforms for the resistance in this circuit.



2. **(a) Circuit** supplied by a steady DC voltage; **(b)** voltage and current waveforms for the capacitor in this circuit.

capacitor and returns to the supply.

We therefore see that the capacitor stores charge, and hence energy, when the magnitude of the current decreases (from time zero to T/4 and from T/2 to 3T/4) either from its maximum positive or maximum negative value to zero.

Power in a pure inductor

The instantaneous voltage, current and power in a pure inductor may be examined in a similar way (*figures 3a* and *3b*). As with the capacitor, the instantaneous power fluctuates giving an average power of zero; the power flows into the inductor when the current increases in magnitude, either positively (from T/4 to T/2) or negatively (from 3T/4 to T).

This can be summarised by noting that when current and voltage are in-phase, power is, on average, consumed by the circuit from the supply; when current and voltage are in phase quadrature, no power, on average, is consumed.

Power in a real circuit

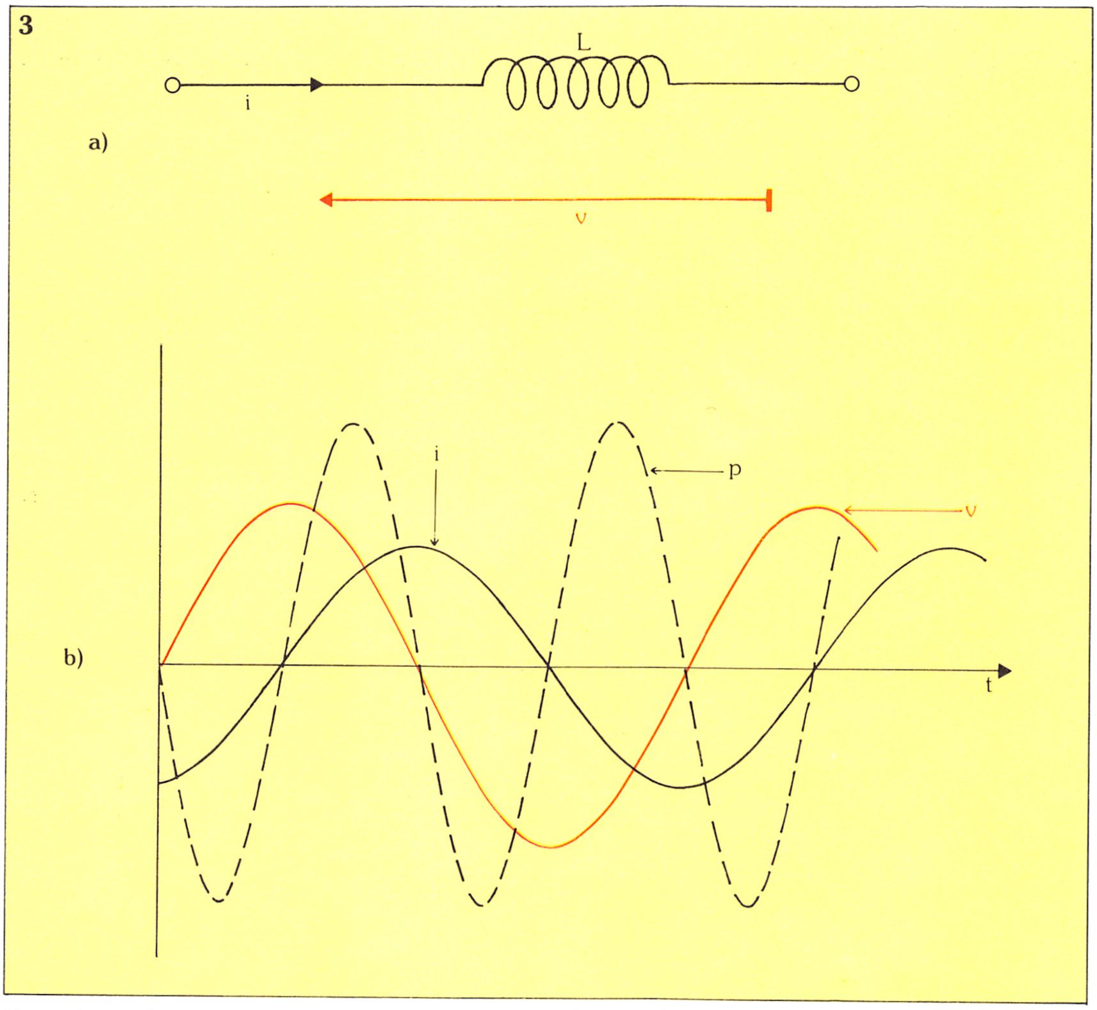
We'll now examine the case of a real circuit – *figure 4a* illustrates an inductor connected in series with a resistor. The instantaneous current and voltage in the circuit are shown in *figure 4b,* where current lags voltage by an angle ϕ. We can see that the instantaneous power obtained from:

$$p = vi$$

is positive, when v and i are either both positive or both negative, and negative, whenever v and i are of opposite polarity. So the power is positive for ωt lying between ϕ and π, and negative for ωt between 0 and ϕ. Since ϕ is never greater than $\pi/2$, the positive power is always greater than the negative power and so, on average, power is provided to the circuit by the voltage supply.

The phasor diagram for this circuit is shown in *figure 4c,* and if we note that the only power dissipated occurs in the resistor, the average power is given by:

$$P = V_R I$$



3. Examining the instantaneous power, voltage and current in a pure inductor.

From *figure 4c* we see that:

$$\cos\phi = \frac{V_R}{V}$$

which leads to:

$$V_R = V\cos\phi$$

Thus we can see that the true power is given by:

$$P = VI\cos\phi$$

which is equal to the heat dissipated in the circuit.

Apparent power

Apparent power is an extremely important measure of the ability of a system to handle current and voltage. When electricity transmission systems are designed, for example, the wires have to be sufficiently large to conduct the required current without overheating, while the insulation has to be strong enough to withstand the maximum voltage. These considerations apply to all electrical equipment – particularly transformers and motors.

We see, therefore, that the product of current and voltage called apparent power, or volt-amperes (represented by the symbol S), is a measure of the maximum voltage and current that an electrical apparatus can withstand. Thus, for the circuit in *figure 4a,* the apparent power is equal to VI.

However, we have seen that the power which the circuit can handle is limited to:

$$P = VI\cos\phi$$

where $\cos\phi$ is always less than 1, giving us a value that is less than the apparent power. The circuit, therefore, has been made to handle power levels greater than those normally required as a safety measure.

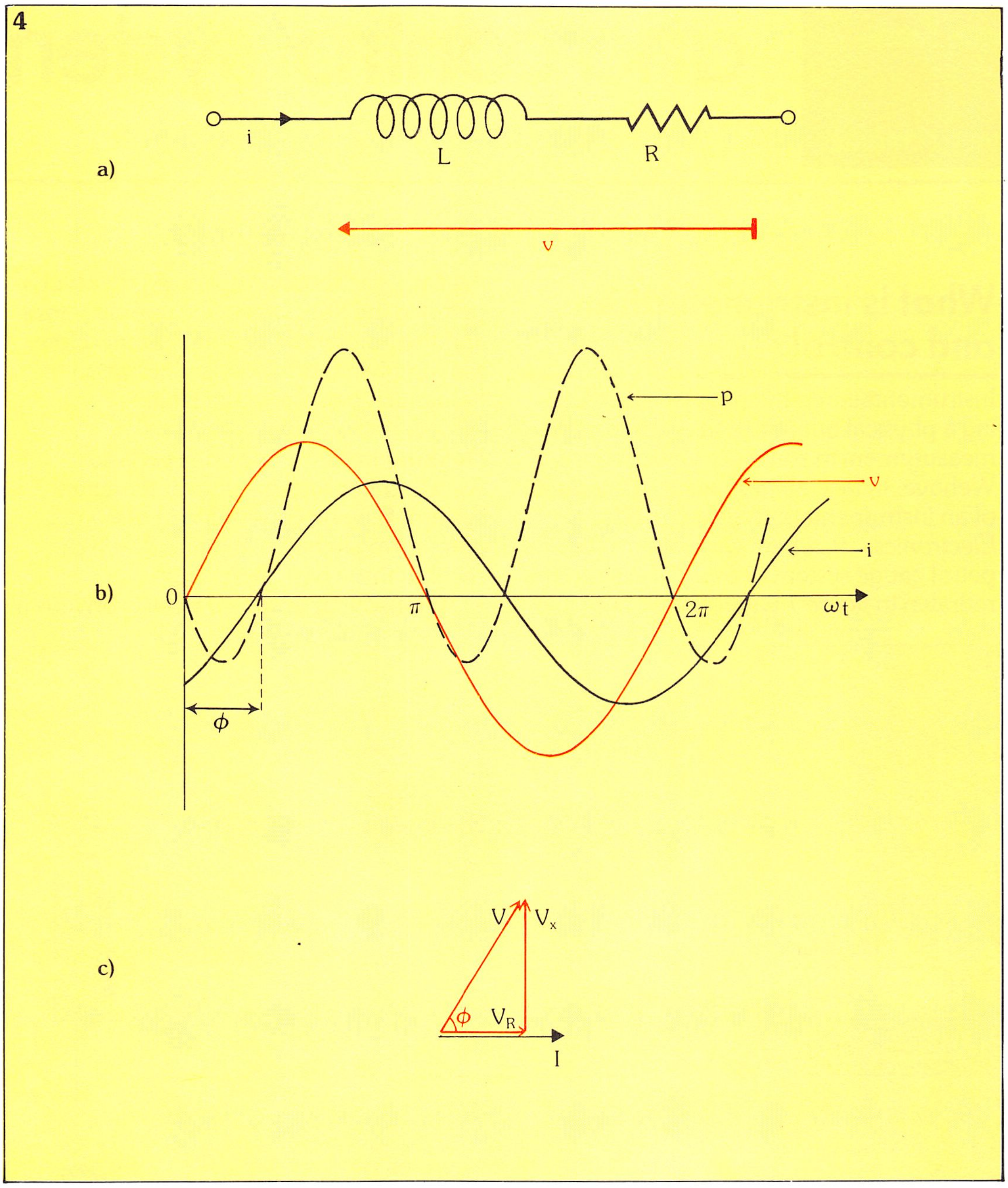
Apparent power is measured in units of volt-amperes (VA) or kilovolt-amperes (kVA).

Power factor

Inefficiency in system construction due to the difference between the real and apparent power ratings is unfortunately inevitable. We shall look at ways in which this can be improved in a later *Basic Theory Refresher.*

The ratio of the real power consumed by a system to its volt-amperes rating is defined as the **power factor**, pf, where:

$$\text{pf} = \frac{P}{S} = \cos\phi$$



4. (a) A real circuit with an inductor connected in series with a resistor; **(b)** instantaneous current and voltage waveforms for this circuit; **(c)** phasor diagram for this circuit.

Reactive power

The **reactive power** of a circuit (also known as reactive volt-amperes) is designated by the symbol Q, where:

$$Q = VI \sin \phi$$
$$= V_X I$$

where V_X is the voltage across the reactive part of the circuit. Reactive power is measured in units of volt-amperes reactive (VAR) or kilo-volt-amperes reactive (kVAR).

Looking at *figure 4c,* a phasor diagram which represents power could be drawn by multiplying all the voltages by the constant current, I. We would see that the apparent power, S, is related to the real power, P, and reactive power, Q, by:

$$S = \sqrt{P^2 + Q^2}$$

and the power factor:

$$\cos \phi = \frac{P}{S}$$

□

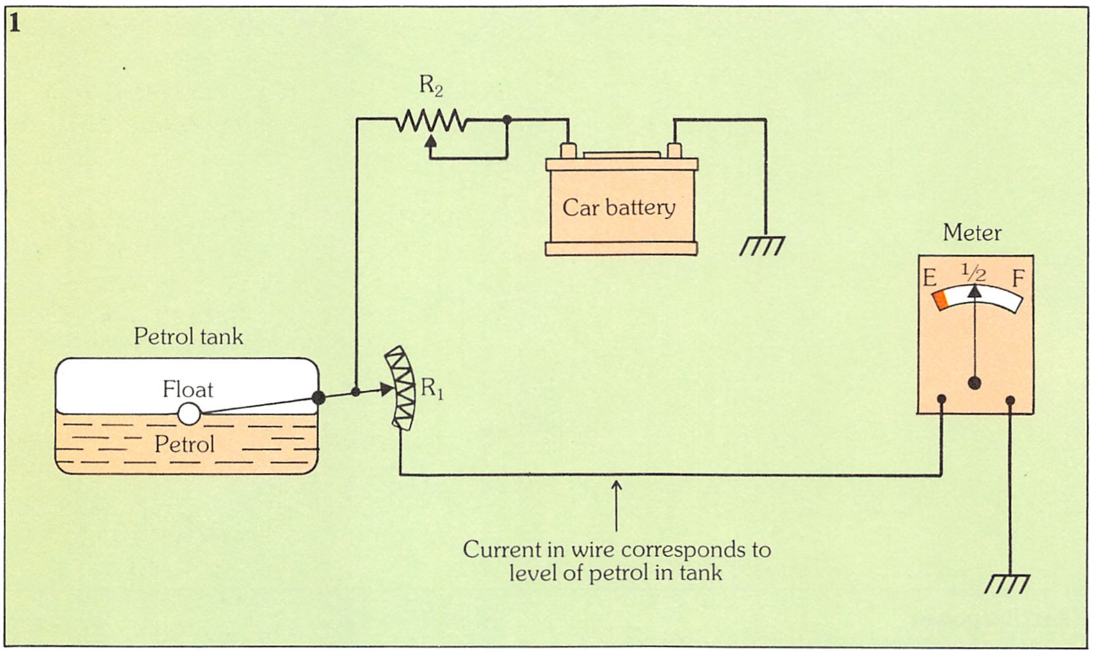
22 DIGITAL ELECTRONICS

Instrumentation and control systems - 1

What is instrumentation and control?

Instrumentation is the process of measuring a physical quantity and reporting that measurement in some convenient form. Wehave, in fact, already seen an example of an instrumentation system in *Digital Electronics – 9,* where we looked at the petrol gauge system of a car. The system is redrawn in *figure 1* where we can see that the level of fuel in the petrol tank determines the height of a float. The float is attached to the end of an arm, which is pivoted so that as the float moves up and down, the other end of the arm acts as the wiper of a variable resistor, R_1. Thus, the amount of petrol in the tank determines the value of the resistance.

The car battery provides a current, through limiting resistor R_2 (a preset resistor) and the variable resistor, which is transmitted along a wire to the ammeter in the car's dashboard display. As the level of petrol in the tank varies, the current varies and the pointer of the ammeter therefore moves, giving the driver an indication of how much petrol is in the tank.



1. A petrol gauge instrumentation system.

Defining instrumentation systems

As with other systems, instrumentation systems may be considered in block diagram form as shown in *figure 2.* We can see that three basic parts exist:

1) A transducer or sensor which converts a measure of the physical quantity into another form. In *figure 1,* the transducer is the float/variable resistor combination which converts a measure of the amount of petrol in the tank to a resistance which controls electrical current. This current is the output signal of the transducer.

2) A signal processor which acts on the transducer's output signal, performs some operation on it, and converts it into a suitable form for display. The signal processor in the petrol gauge example is the

limiting resistor, R_2, which prevents too much current flowing through the ammeter.
3) A display device which converts the signal at the output of the signal processor into some readable quantity. The ammeter, of course, is the display device of our example instrumentation system.

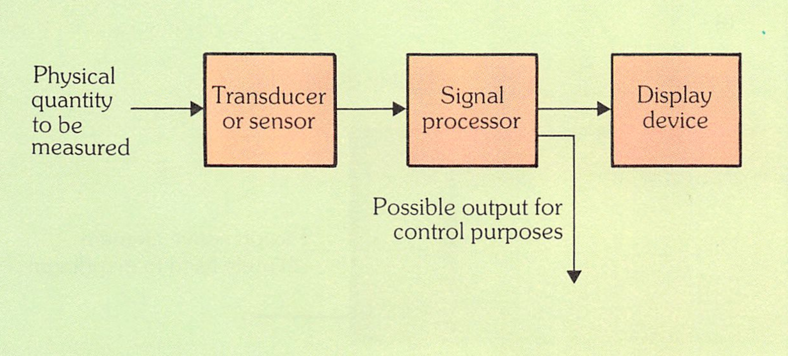
Also shown in the block diagram of *figure 2* is an output, after signal processing, which is used for possible control purposes. Complete instrumentation systems are often used as part of a control system, to provide measurements of some quantity or quantities to be controlled. We shall be looking at control systems a little later.

Electronic instrumentation

The block diagram shown in *figure 2* is true for all instrumentation systems. The petrol gauge example is an electrical instrumentation system in which petrol level is turned into electric current (with an electromechanical transducer). Instrumentation systems may also be totally electronic, with no moving parts and have active, amplifying devices to process the transducer's output signal. Others may be purely mechanical; purely pneumatic (e.g. working on the pressure of gas, or a vacuum); or a combination of two or more methods.

Our study of instrumentation systems will, however, be restricted to electrical/ electronic methods and, in particular, we shall look at further examples used in cars.

2



2. Block diagram of an instrumentation system.

Examples of car instrumentation systems

Instrumentation systems in cars include the equipment and devices which measure vehicle variables and engine parameters, and display their status to the driver, in some way. Examples include:
1) speedometer;
2) odometer, i.e. the distance meter;
3) tachometer, i.e. engine rev meter;
4) fuel gauge;
5) oil pressure gauge;
6) clocks;
7) door-ajar indicators;
8) ammeters;
9) voltmeter.

Many of these systems are non-electronic, however, manufacturers are increasingly becoming aware of the advantages of electronic instrumentation, and the previous slow trend towards the use of electronics is now beginning to speed up, with many of the newest car models factory-fitted with a number of electronic instrumentation systems.

Whichever type is used, however, the general goal of the instrumentation system is universally the same – to provide an accurate, reliable and readable measurement of a particular physical quantity.

Accuracy

The accuracy of an instrument, i.e. how close the instrument's measurement is to the actual physical quantity, is generally specified as **error**, which is expressed as the difference between the actual and measured values. For example, we might look at a clock and read the time as a quarter past one; the actual time might be fourteen minutes past one. In this case we would say that the error was one minute.

In instrumentation systems, differences between actual and measured values arise from two important types of error – systematic error and random error.

Systematic errors are due to such things as improper calibration, drift caused by temperature variations, or improper use of the instrument. Systematic errors are predictable and can be corrected for, or eliminated by, the user.

Random errors, on the other hand, are caused by electrical, mechanical, or any other type of noise or disturbance which adversely affects the measurement of the physical quantity. These errors can be minimised by careful design, and the addition of special noise reducing circuits.

Factors affecting accuracy

A number of system factors can affect the accuracy of an instrumentation system. Many of these, if left unchecked, may produce systematic errors which degrade overall performance. The first factor we shall consider is the time taken for the output of the instrumentation system to change in response to a change of input – **response time.**

Generally, an instrumentation system's response time should be quite fast, so that the human observer sees a change in the physical quantity being measured as soon as possible. The driver of a car, for example, should be quickly informed of a loss or increase in oil pressure, so that the engine may be immediately stopped, preventing damage. An example of an instrumentation system which could be used in this application is shown in *figure 3a,* where a pressure transducer is shown connected to a regulated voltage, of 10 V DC, and a series resistor. The resistance of the transducer varies with engine oil pressure so the transducer/series resistor combination is effectively a variable potential divider, the output voltage of which is dependent on engine oil pressure.

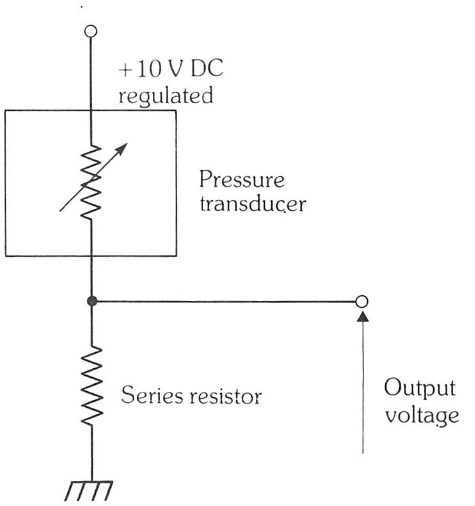
Figure 3b illustrates the basis of a suitable transducer for this application which uses a **piezoresistive element** to detect pressure. **Piezoresistivity** is a phenomenon which occurs in certain semiconductors where the actual resistance changes in proportion to the **strain** (fractional change in length). As the piezoresistive element is securely attached to a flexible diaphragm, which bows out as engine oil pressure increases, the length of the element also increases, as shown in *figure 3c.*

In some cases, though, a fast response time is undesirable – the petrol gauge is one such example. As the car moves, so does the petrol in the tank, and the float on the surface of the petrol therefore moves rapidly up and down. If the response time of the system was fast, the pointer on the ammeter would also move rapidly up and down, and the driver would not be able to take an accurate reading of petrol level.

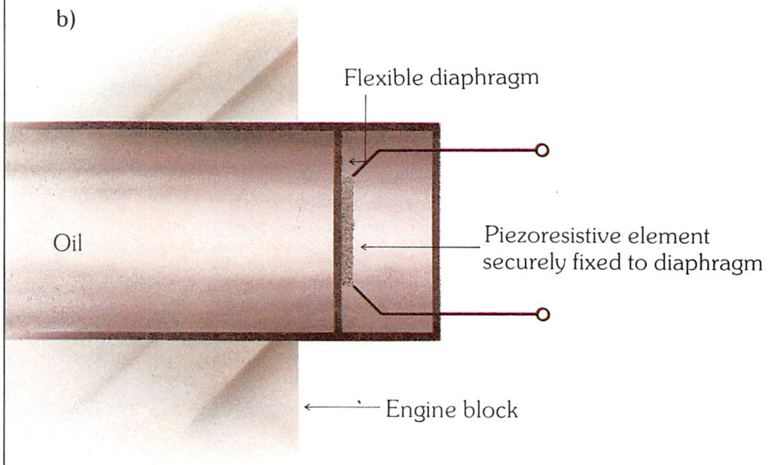
One way of reducing the response time of such an instrumentation system is

3

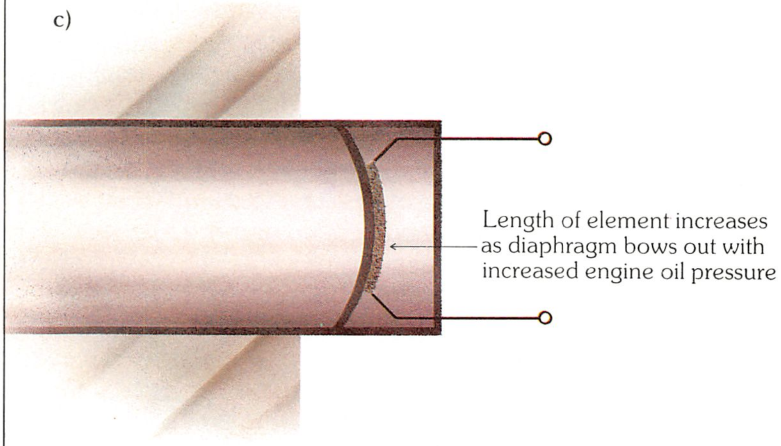
a)



b)

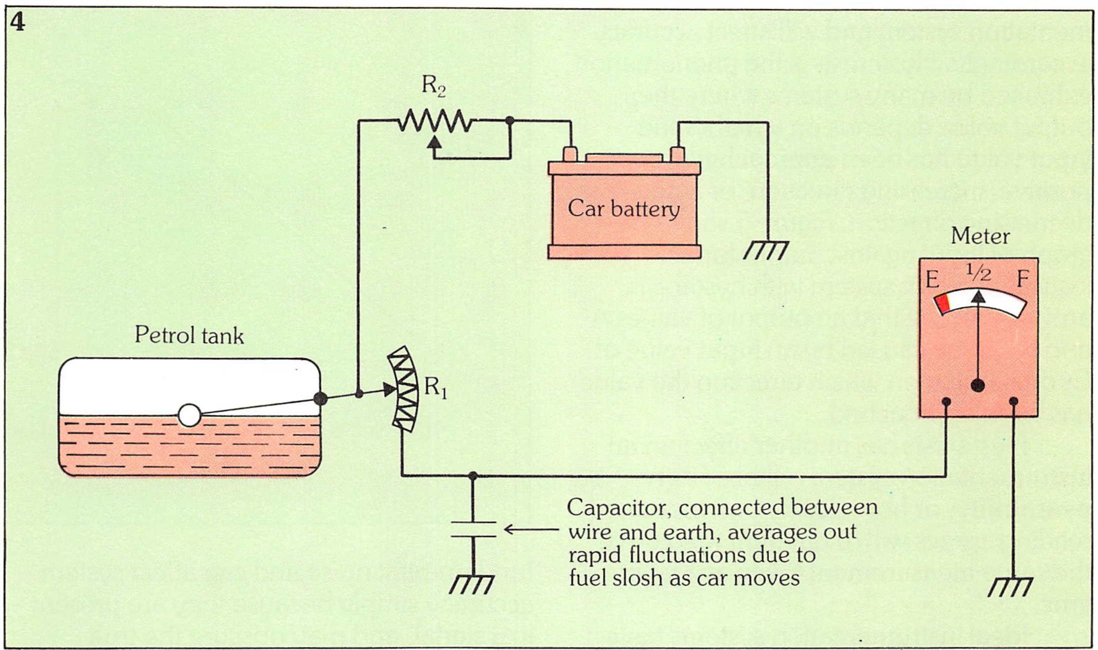


c)



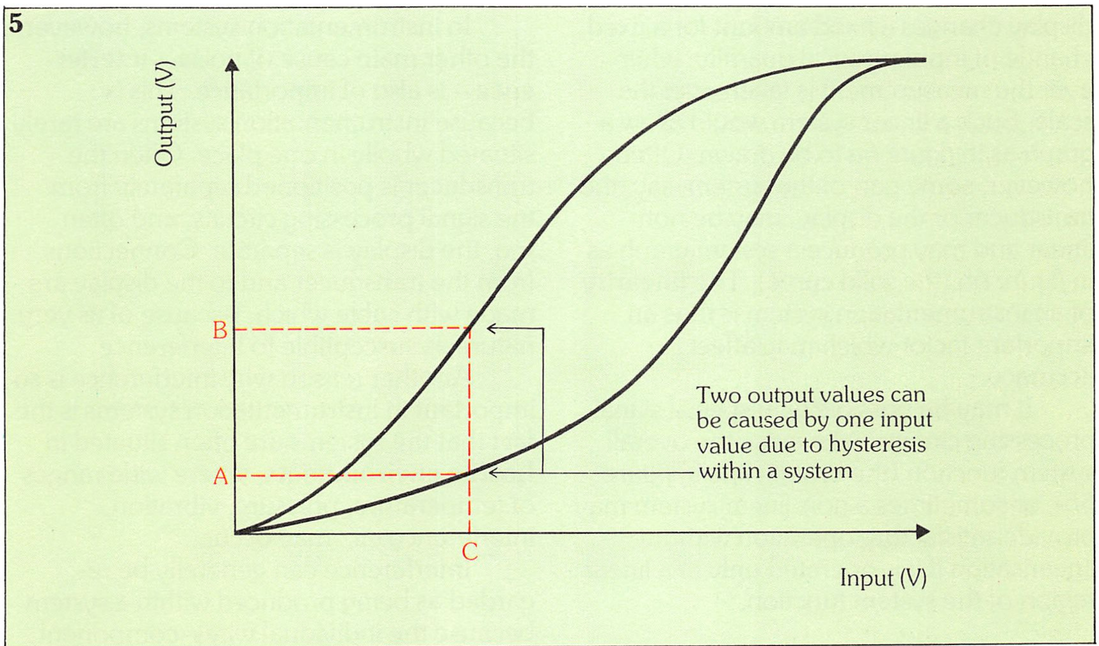
3. (a) A pressure transducer used to detect increase or loss in oil pressure; **(b)** the piezoresistive element detects pressure; and **(c)** lengthens as the oil pressure increases.

4



4. When response times need to be reduced, a capacitor can be used to average out fluctuations in current.

5



5. Input *vs* output for an instrumentation system with hysteresis.

shown in *figure 4,* where a capacitor is connected between the variable resistor transducer and earth. The effect of the capacitor is to average out the rapid increases and decreases of current flowing from the variable resistor, so that the ammeter reads an average value, more closely reflecting the actual petrol level.

Closely associated to the response time of a system is its **bandwidth**, defined in *Solid State Electronics 24* as the range of frequencies over which the response does not fall by more than 3 dB from the mid-range value. If the bandwidth of a system does not extend to a high enough frequency, rapid fluctuations of the physical quantity measured by the transducer may not be passed through the system, and so may not be displayed quickly enough. Conversely, by reducing the upper frequency limit of the frequency range bandwidth, the response time of a system is increased – this is, in effect, what occurs when the capacitor is added to the petrol gauge circuit.

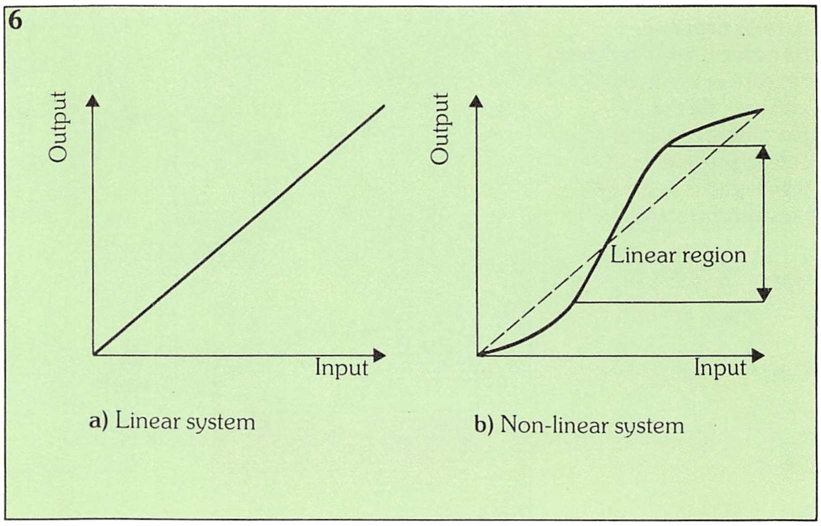
Hysteresis may occur in an instru-

mentation system and will affect accuracy accordingly. Hysteresis is the phenomenon exhibited by many systems where the output value depends on whether the input value has been approached from a positive, increasing direction, or a negative, decreasing direction. *Figure 5* shows a graph of input against output for an instrumentation system with hysteresis, and we can see that an output of values A *and* B can be caused by an input value of C, depending on which direction the value has been approached.

Hysteresis has another effect on an instrumentation system – that of its **repeatability**, or how closely one measured reading agrees with a different reading of the same measurement taken at a later time.

Ideal instrumentation systems have overall linear functions, i.e. the output display changes a fixed amount for a fixed change of input physical quantity, wherever the measurement is taken over the scale. Such a linear system would allow a graph as in *figure 6a* to be drawn. Often, however, some part of the system, say, the transducer or the display, may be non-linear and may produce a system graph as in *figure 6b* (the solid curve). The **linearity** of an instrumentation system is thus an important factor which may affect accuracy.

It may be possible with special signal processing circuits to linearise the overall system function (the broken line in *figure 6b*), or sometimes a non-linear system may provide satisfactory operation without linearisation if it is operated only in a linear region of the system function.



6. Input *vs* output for: (a) linear; and **(b)** non-linear systems.

Random error factors

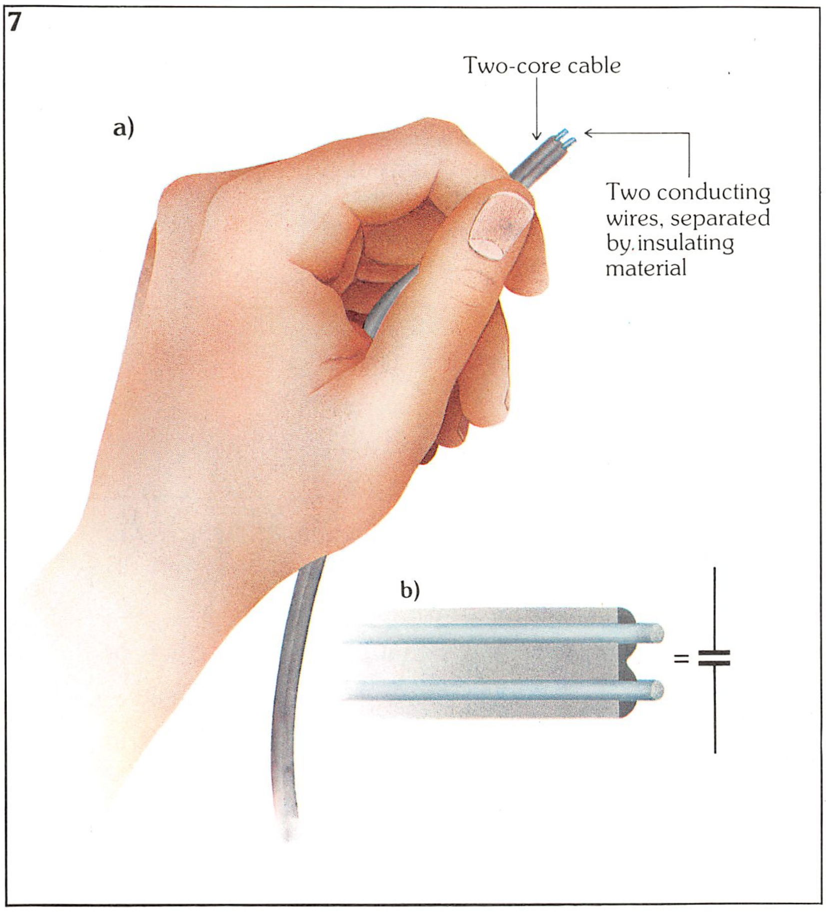
All of the factors affecting system accuracy that we have so far seen have been systematic errors which can be eleiminated by careful system design and operation. There is, however, one main group of factors affecting accuracy which are random errors and cannot be eliminated. The group is classed under the title of **noise**. In *Solid State Electronics 15* we looked at the fundamental, random noise which occurs in the electronic components of any circuit. Thermal noise, shot noise and flicker noise are the three most important types of fundamental noise and can affect system accuracy simply because they are present in a signal, and may obscure the true measured signal, thus giving false readings.

In instrumentation systems, however, the other main cause of noise – **interference** – is also of importance. This is because instrumentation systems are rarely situated wholly in one place. Often the transducer is positioned separately from the signal processing circuits, and often too, the display is separate. Connections from the transducer and to the display are made with cable which, because of its very nature, is susceptible to interference.

Another reason why interference is so important in instrumentation systems is the fact that the systems are often situated in **hostile environments**, where wide ranges of temperature, pressure, vibration, interference etc. may occur.

Interference can generally be regarded as being produced within a system because the individual wires, component leads, cables etc. effectively act as components. For example, the length of two core cable shown in *figure 7a* consists of two conducting wires, separated by an insulator, and is thus similar in construction to a capacitor (*figure 7b*). It two lengths of cable are positioned side by side, further capacitances are formed between the two cables, and any signal on one cable is partly passed, via the capacitances, to the other cable. Such interference is known as **capacitive interference.**

The other main form of interference is shown in *figure 8* where a transducer is



AC circuit or wire passes, inducing an AC voltage of the same frequency in the loop. This type of interference is known as **inductive interference.**

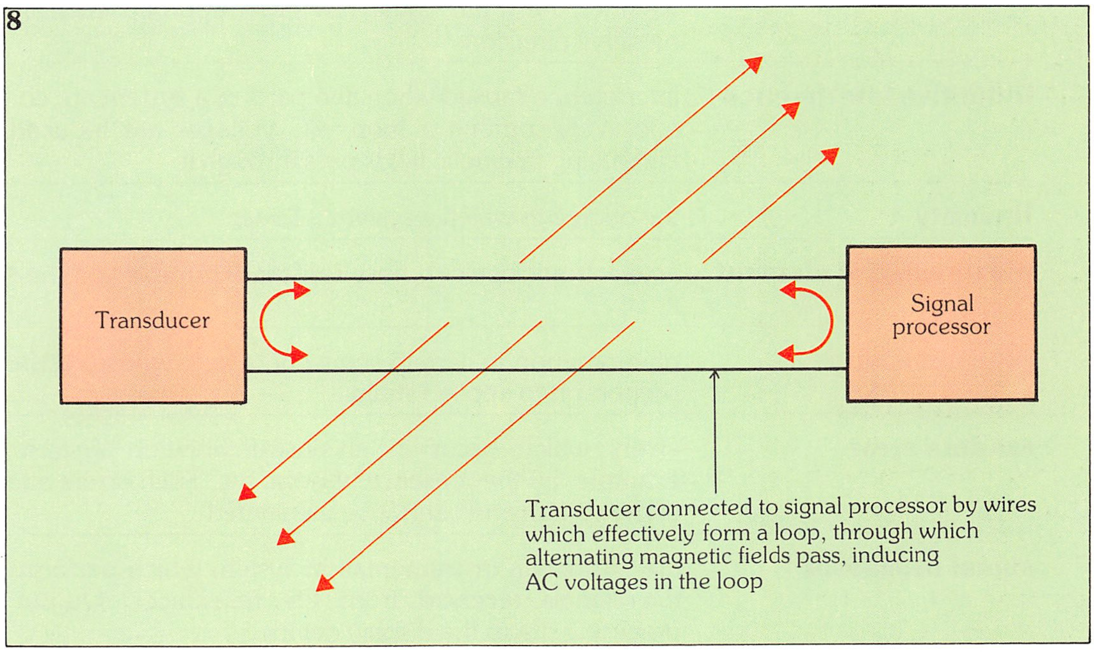
Two main causes of capacitive and inductive interference are common: **hum**, caused by close proximity of power lines, or power circuits through which mains frequency currents pass; and **cross-talk** from signal lines or circuits. Hum produces interference noise of the same frequency as mains, i.e. 50 Hz, but cross-talk interference frequency depends on the frequency of the signals in the interfering lines or circuits.

Reliability and readability

An instrumentation system's reliability refers to its ability to perform its function, accurately and continually, whenever required, under specified environmental conditions, for an adequate period of time. Reliability must, therefore, be built into a system by using reasonable design margins and components which will operate over the temperature range and environmental conditions in which the system is likely to operate.

7. Two-core cable (a) is similar in construction to the capacitor; **(b)** producing capacitive interference.

8. Inductive interference.



connected by a two core cable to the signal processing circuit of an instrumentation system. The effect of this connection is to create a loop of wire, through which the alternating magnetic field from a nearby

The readability of an instrument describes how easily the results of a measurement can be read and understood by a user. A properly displayed measurement should indicate the numerical value of the

measurement, along with the units of the measured quantity – this is called **calibration**. For example, on a speedometer indicator dial, not only should the major divisions be marked with miles per hour and kilometres per hour, but also subdivisions that make sense to the user at a glance.

Readability also depends on the physical aspects of the display, such as size, position, the angle of view, how well it is lit, and the contrast of the display elements to their background.

In *Digital Electronics 23* we shall turn our attention to the individual types of transducers available for use in instrumentation and control systems and also look at some applications.

Glossary

calibration	process of displaying a numerical measurement along with the units of the physical quantity measured
capacitive interference	interference between two parts of adjacent circuits, in which the individual wires, components leads, cables etc. effectively act as capacitors
cross-talk	interference caused when signals in one circuit pass by capacitive or inductive means to another circuit
hum	interference to a circuit, caused when power lines or circuits are in close proximity. Hum may be caused by capacitive or inductive interference, and the interference frequency is that of the mains, i.e. 50 Hz
hysteresis	phenomenon which causes a system's output value to depend on whether the input value has been approached from a positive, or negative direction
inductive interference	interference caused when two parts of a system are connected by wire effectively forming a loop. AC voltages are induced in the loop if alternating magnetic fields pass through it
linearity	the degree to which a system is linear
piezoresistive element	an electronic device whose resistance depends on the strain applied to it
piezoresisivity	phenomenon in certain semiconductors, whose resistance changes in proportion to applied strain
random error	errors which occur in an instrumentation system produced, for example, by the presence of noise etc. Such errors affect the accuracy of the system and cannot be eliminated
signal processor	the part of an instrumentation system which performs operations on the signals received from the transducer. Output of the signal processor drives the display device
systematic error	an error in an instrumentation system caused by, say, improper calibration, improper use etc. Systematic errors are predictable and can be corrected

1

COMPUTERS & SOCIETY

Technology in the home

Introduction

In this series of articles we will be exploring the different uses and functions of the 'new technologies'. We will also examine their role in society today and speculate a little on how they might shape the future.

The word 'technology' is defined by Chambers as: 'the practice of any or all of the applied sciences that have practical value and/or industrial use: technical method(s) in a particular field of industry of art'. The concept applies equally to the simplest of hand tools or the latest miniaturised marvels: both have been created and developed to perform specific functions useful to society.

The relationship between technological

Below: Prestel's new Micronet 800 service allows home computer uses access to over 2,000 programs free of charge.

The Research House/British Telecom

development and societal change, however, is not so clearly defined and is extremely complex. Will a society of increasing sophistication demand technological development to further refine its life-style, or are technological developments simply picked up and applied by society as it thinks fit?

Although we are now living in a post-industrial society, the effects of technological change are no less great than those that took place during the industrial revolution. In Britain, Western Europe and the U.S.A. predominantly agricultural economies and societies were transformed into industrial societies. Capital, raw materials, labour, expanding markets and improving transport all came together to create an atmosphere in which technological advances kept on being made, improving efficiency and profits.

The textile industry was the first to experience this sort of change as the flying shuttle (1733) and the power loom (1785) replaced the old hand-looms. The increasing mechanisation of agriculture, the enclosure of arable land and the desire to seek a better standard of living, led people away from their old rural communities into the growing cities. Today, however, these large industrial conurbations are breaking down. Industrial (and office) automation is releasing large numbers of people from boring, repetitive tasks – but to do what? Our new technologies present a social challenge to which all of us, not only scientists, governments and economists, should address ourselves.

Changing technology in the home

The home is usually the last place to be affected by technological development, unless those developments are specifically aimed at the domestic market. In the past, improvements in productivity and profitability in the home were not seen as important considerations – one could always employ more servants! It is only in recent times, with the decline of the servant class, improvements in the general standard of living and the emergence of the family (especially those staying at home) as a 'consumer unit', that domestic technology has assumed such economic importance.

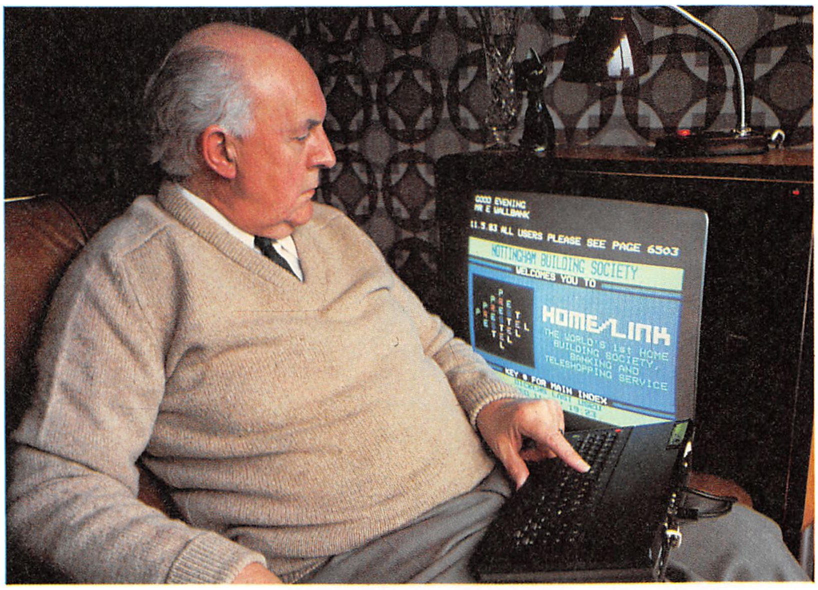
Over the last hundred or so years, developments in electrical and electronic technology have undergone considerable acceleration until they are now almost exponential. The telephone was patented by Alexander Graham Bell in 1876 and was the first electrical device to enable the two-way transmission of speech. Previously, the only 'instant' long range communication was the morse telegraph – only really effective in the hands of a skilled operator. With Bell's invention, instant two-way communication was available to everyone; the telephone was adopted first by business and commerce, and later by those private individuals who could afford it. Since then, of course, telephones in the home have become the rule rather than the exception.

The second electronic device to find acceptance in the home was also a com-

Left: a home computer connected to a modem.

munications tool, albeit only one-way: the radio. Although the first radio sets were, by and large, constructed by home enthusiasts (following the plans published in the predecessors of today's hobbyist electronics magazines), radio broadcasting by the British Broadcasting Company was set-up as a commercial enterprise, making its money from the sale of receivers. To achieve this end, the programmes had to possess a certain mass public appeal to induce large numbers of people to take their entertain-

Below: Prestel 'Homelink' will enable banking and shopping from home at the touch of a button.



ment *in the home*. Thus, broadcasting began to compete with the more traditional entertainments, cinemas, theatres and sporting events, which people had to leave home to see. Information was now brought into the home with an immediacy hitherto unknown.

The radio and the telephone hold the key to the development of many of today's communications systems: television was developed as a form of visual radio transmission. The television and the telephone still provide the primary vehicles whereby information and entertainment are brought into the home; although the nature of that information and entertainment is changing.

The changing television set

Looking at the present domestic applications of electronic technology we see that electronic devices in the home broadly fall into two categories: 'labour-saving' service devices, and entertainment and/or information providers.

The domestic entertainments industry is burgeoning, video recorders being but one example. As you probably know, these machines can be used in three different ways: to record broadcast programmes, to be replayed when it suits the viewer; to record 'home videos' with the addition of a lightweight camera; and to view the large range of material available to rent or buy on pre-recorded tapes.

Cable and satellite television will offer considerably wider entertainment options. These services will sell programmes to subscribing viewers on a 'pay according to what you receive' basis. The viewer will have to pay for any specialised equipment, such as the buying or hiring of a receiving dish for satellite TV, and the cabling and alterations to existing receivers for cable TV. The limiting factor in the installation of a cable TV network will be the cost of installing the copper and fibre optics needed for the network.

The cables along which cable TV programmes are transmitted will become an important part of the domestic electronics of the future. Existing cable television services (most of which were set up in the 1950s and 60s to relay the broadcast channels to areas that could not receive RF signals) use either ordinary coaxial or multi-stranded cables as the transmission medium. These types of cable are limited in the number of different signals they can carry.

Fibre optic cables, however, are able to carry a very large number of signals and also possess many other advantages. The cable consists of one or many thin strands of glass fibre which has excellent light transmission characteristics. A laser beam can be modulated to transmit digital information in the form of pulses of light. Many separate signals can be multiplexed (mixed) together in one fibre optic strand and the overall information carrying capacity is many times greater than conventional cable.

Television receivers can also be adapted to receive the broadcast **teletext** services. Teletext is a computerised system that broadcasts digital information in unused lines of television pictures. This information is decoded within the receiver by special circuits and presented as pages of text and diagrams. Originally developed for providing programme subtitling for the hard of

hearing and the deaf, teletext is a free domestic service also providing news updates, weather forecasts and travel information.

Viewdata is a similar text transmission system that utilises a special television adaptor and a telephone link to a central data base, for example, British Telecom's Prestel system. Unlike teletext, viewdata is essentially business oriented, the information provided being paid for by the consumer. Such things as stock market quotations,

The telephone link

Telephones can now be used to transmit data around the world – direct from computer to computer. This type of data link can be employed equally by scientists or businesses on either side of the Atlantic, or by the travelling salesman and his head office within the British Isles or indeed one city. All the travelling salesman needs is an acoustic coupler or modem attached to his portable computer. The telephone handset in the hotel or home is connected to this,

The Research House/British Telecom

Left: 'Picture Prestel' is hoped to be in operation in about 12 months time. The bandwidth of telephone lines is not wide enough to send pictures, however, by sending portions of the picture at a time, so that the picture builds up slowly, BT hope to overcome this.

sports results, and travel information are available.

The subscriber is able to interact with the information provider's computer in some cases, and private viewdata systems can also be implemented. Businesses use viewdata to transmit information to offices or branches around the country. It is also being used as an **electronic letterbox**, whereby data can be 'posted' to an individual's home, or hotel.

The television receiver can also be used as the screen end of many home computer terminals. The BBC micro and the Sinclair ZX81 are examples of such machines. TV gaming consoles, connected to the TV receiver, are also becoming increasingly popular.

and data can be received or sent.

Using microprocessor technology, it is now possible to program telephones with a list of numbers – the phone then 'keeps trying' these numbers until a connection is made.

Other digital devices in the home

Music has also been touched by the electronic revolution, both in its performance and reproduction. People once played acoustic instruments and listened to wind-up gramophones, but now electronic music is widespread and quite advanced, even in the home, where small synthesisers like the Casiotone are cheap and popular instruments. With continual improvements in equipment and component design, the

Right: British Telecom's electronic 'Ambassador' phone. The microphone facility on the right enables the user to speak from a distance.

The Research House/British Telecom

reasonably cheap hif-fi systems bought today produce a quality of reproduction which is considerably better than the more expensive systems (even allowing for inflation) bought ten years ago.

Developments in digital electronics are now changing the nature of sound broadcasting and recording. Digital recording equipment is now being used more and more, the BBC now uses digital relay and recording systems in its broadcasting and recording of concerts and music programmes. The digital techniques being developed now mean that the recording and creation of music, drama and other sound programmes will become more flexible and advanced. Compact disk (CD) players are now available which are digital optical devices, read by modulation of incident laser light.

Electronic and digital systems can be used to control a home's domestic functions. It is possible to automatically regulate temperature, humidity and lighting while the entire building and grounds can be electronically monitored for fire, and intrusions, with anything amiss being immediately relayed to a control point or police station that may be miles away. Such systems are very expensive due to the cost of transducers.

Domestic appliances, such as washing machines, are now controlled electronically, with microprocessors considerably simplifying the circuitry previously required by conventional electromechanical program controllers.

The future

The fiction of the fifties proclaimed the brave new 'future world' of the eighties, but we are still driving petrol engined cars and our homes are not tended by robot servants.

Although great leaps forward have been made in scientific achievements, their application to the technology of everyday life has not effected dramatic change. However, looking into the not too distant future, it is possible even with the technology available today, that the home is likely to become a rather different place.

The changes that will occur during the next twenty or so years will be gradual but no less radical or lasting than those changes during the industrial revolution. The movement of people away from their homes as places of work during that time is now being reversed, as small home-based cottage industries are being actively encouraged and the concept of the home as being an extension of the computer-based office is now a reality.

The Reseach House/British Telecom

Left: a new telephone incorporating an LCD display.

The term **information technology** refers to the field of creating systems and devices used for the storage and dissemination of information. Most of the devices and applications we have discussed here (and will go on to discuss in greater detail in later chapters of this series and the *Communications* series) have done exactly that, by means of audible, visual or digital signals. More importantly, all of these devices and applications are capable of being made to work with digital signals if they don't at the moment. The key to the future lies in the way that these autonomous and separate devices can be *linked together.*

If every home had a terminal or terminals comprising devices for data capture, display, manipulation and transmission, data interaction could be channelled through a central data exchange or exchanges in the manner of a conventional telephone exchange. This can now be made possible with fibre optic technology. You'll recall that we mentioned the superior information transmission characteristics of fibre optics earlier. If each home in the country were to be equipped with a fibre optic link-up, say, for the supposed primary purpose of supplying cable television, then the whole country could be economically linked to central data exchange points – both private and public.

As well as supplying television pictures and computerised information services, this system opens up a whole new range of possibilities. As it is possible to multiplex (mix together) many separate digital signals, the fibre optic cables could carry a variety of different signals to different places. Telephone conversations could be linked side by side with such functions as automatic gas and electricity meter reading, while new services, such as home banking and shopping by television (under trial at the moment), would be possible. This concept of a completely computerised and linked home is known as the **hard-wired house**.

There are, of course, many social aspects involved in the concept of the hard-wired house. Taken to its extreme, families might become isolated units in the home – with parents working at home and any children receiving their schooling at home also. However, this may lead to a redevelopment of smaller communities, social contacts no longer being provided by the workplace.

The home is now being increasingly seen as a target market by large and small 'high technology' companies alike, with 'technology' now being specifically developed for it.

25

SOLID STATE ELECTRONICS

Operational amplifiers - 2

In *Solid State Electronics 24* we saw how feedback can be used to change the gain of an op-amp: from its open loop gain, A_{OL}, to the closed loop gain, G.

With purely resistive negative feedback, the bandwidth of the op-amp increases as the gain decreases. So, if the gain of the op-amp is reduced with the use of feedback by, say, a factor of 20, the bandwidth increases by the same factor.

We looked at two main groups of amplifiers: the inverting amplifier, shown in *figure 1,* where the input is to the inverting input of the op-amp; and the non-inverting amplifier, where the input signal is applied to the non-inverting op-amp input, shown in *figure 2.* Although the circuits of the two amplifiers have been drawn in a slightly different way those previously seen, we should be able to follow their operation.

Closed loop gain of any amplifier is:

$$G = \frac{V_{out}}{V_{in}}$$

and for the inverting amplifier this equals:

$$G = -\frac{R_2}{R_1}$$

if the open loop gain, A_{OL}, is sufficiently large. Closed loop gain of the inverting amplifier is similarly given by:

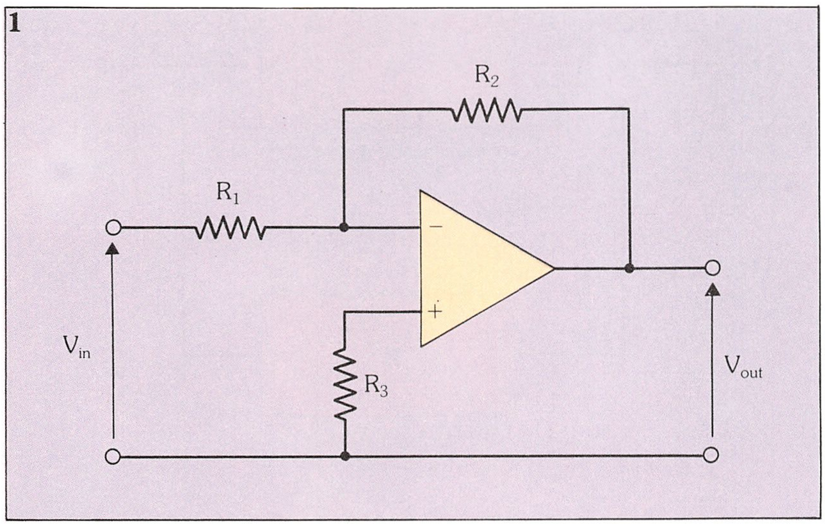
$$G = \frac{R_1 + R_2}{R_2}$$

In both cases, closed loop gain has been approximated by the ratios of only two resistances. This is a very significant factor when dealing with operational amplifiers, because it means that an amplifier's gain and bandwidth can be specified with reasonable accuracy by the designer.

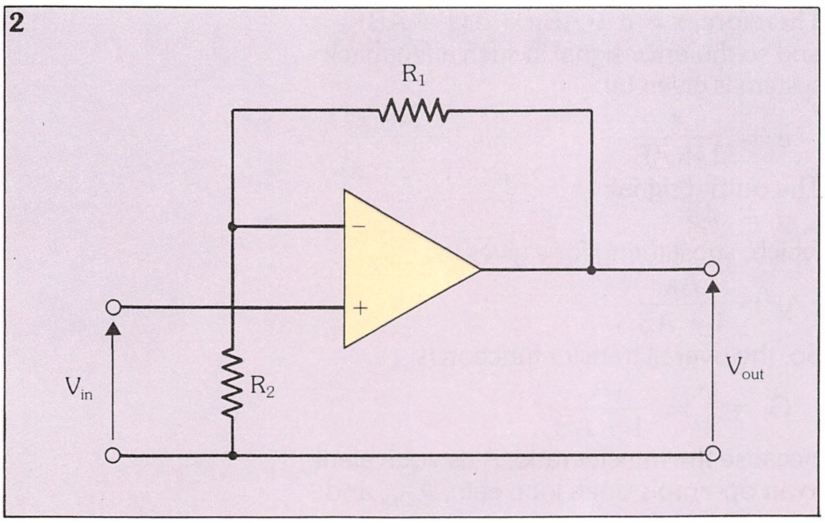
The components used to define the op-amp's closed loop gain (resistors, in these examples) are passive components whose values do not vary to any significant degree as outside factors, such as temperature, change. An op-amp's open loop gain, on the other hand, is highly dependent on temperature and supply voltage for example, so by adding negative feedback we have introduced stability into an otherwise potentially unstable system.

An amplifier's gain and bandwidth are only two parameters which a circuit designer needs to define. Other parameters, also of great importance, are input and output impedances, noise and distortion. The question we now need to ask is what effect, if any, does feedback have on these parameters?

1. The inverting amplifier.



2. The non-inverting amplifier.



Looking at feedback another way

An alternative way of looking at a negative feedback system is shown in *figure 3*. This type of diagram can be used, not only in op-amp circuits but also in other types of feedback systems. We'll see examples of other feedback systems in *Digital Electronics 23* when we look more closely at control systems.

In *figure 3*, the op-amp is shown simply as a block with a voltage gain of A. Because the output of this block is A times the input, so that the output and input are related by a ratio, a block's gain is often known as its **transfer ratio.** In general, the transfer ratio, of any circuit, sometimes known as the **transfer function,** is a mathematical expression which relates the circuit's output to its input.

The feedback circuit of the op-amp is shown in *figure 3* as a block with a transfer ratio of B, where B is equivalent to the feedback fraction, β. The final part of the block diagram is the circulation symbol called a **summing point,** which is a notational way of describing the fact that the feedback circuit allows a certain fraction of the output signal to be fed back to the input of the system and added to the input signal. The fact that the feedback signal is actually *subtracted* from the input signal – shown by the minus sign in the bottom quadrant of the summing point symbol – tells us that the system has negative feedback. A postive feedback system would have a plus sign in that quadrant.

In order to use block diagrams such as this to model feedback systems, we must make two assumptions: first, each block diagram within the diagram may only represent a linear circuit or part-circuit; second, each circuit has infinite input impedance and zero output impedance so that blocks may be interconnected without loading effects. For example, the feedback circuit must not load the op-amp by drawing any current from it, otherwise the transfer ratio of the loaded block is incorrect and needs to be recalculated.

A final point to remember when using the block diagram notation to describe feedback systems is that all signals entering or leaving a summing point must have the same units. In the case of our op-amp block diagram, all signals in the diagram have voltage units, but other systems may have other signals with different units.

The output of the summing point is an **error signal,** e, which is multiplied by the op-amp transfer ratio A, to give the output signal, y, where:

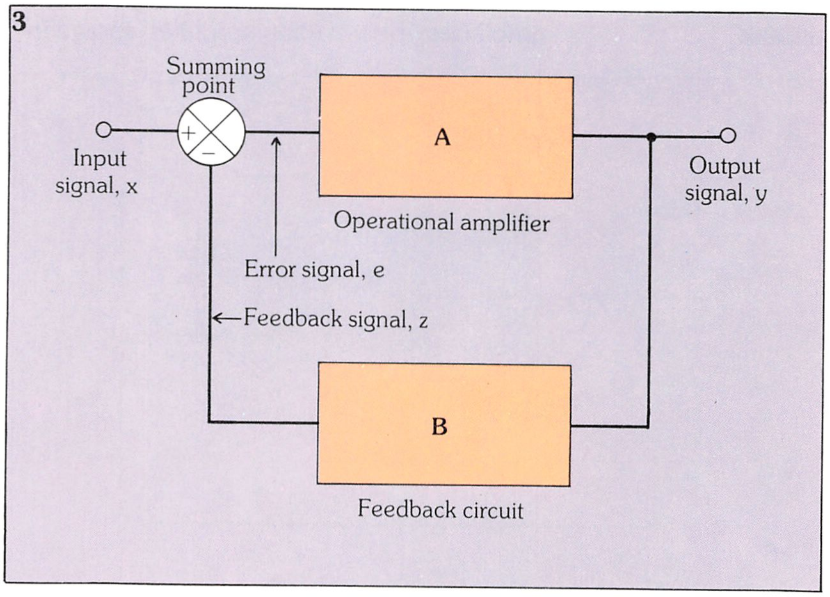
$$y = Ae$$

Now, the feedback signal:

$$z = By$$

But, the error signal, e, is the input signal minus the feedback signal, so:

$$\begin{aligned} e &= x - z \\ &= x - By \\ &= x - ABe \end{aligned}$$



3. Representing a negative feedback system in block diagram form.

Therefore, $x = e + ABe = e(1 + AB)$, and so the error signal in such a feedback system is given by:

$$e = \frac{x}{1 + AB}$$

The output signal:

$$y = Ae$$

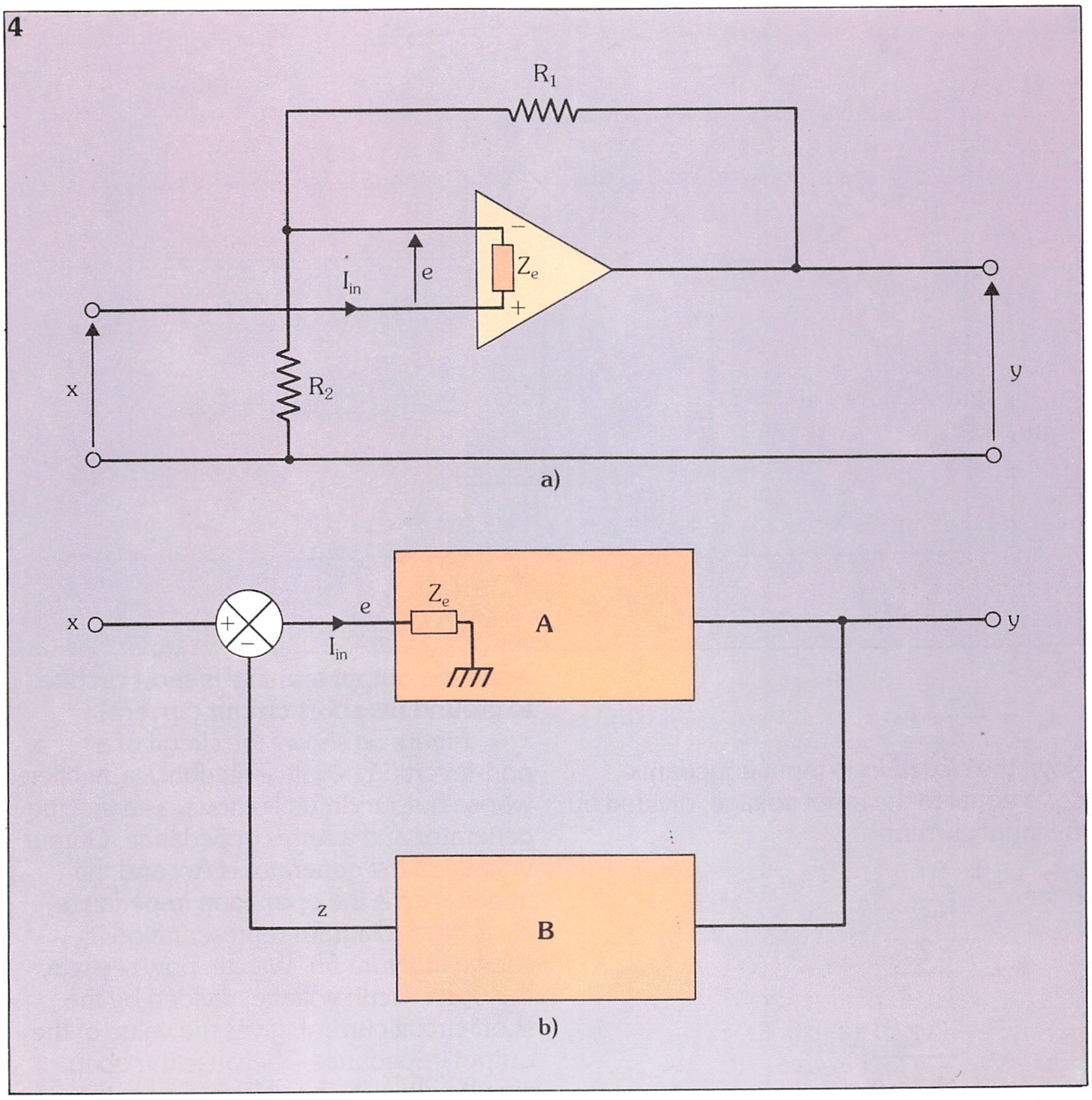
which, substituting for e gives us:

$$y = \frac{Ax}{1 + AB}$$

So, the overall transfer function is:

$$G = \frac{y}{x} = \frac{A}{1 + AB}$$

Because the transfer ratio, A, is equivalent to an op-amp's open loop gain, A_{OL}, and



4. (a) A non-inverting amplifier circuit; **(b)** its block diagram equivalent.

the feedback circuit transfer ratio, B, is equivalent to the feedback fraction, B, this equation is of the same form as the gain formula seen in *Solid State Electronics 24.*

The quantity AB of the diagram, is known as the **loop gain** or **return ratio** and, for large values of loop gain, the actual gain of the whole circuit can be approximated by:

$$G = \frac{A}{AB} = \frac{1}{B}$$

Thus, we can see that variations of open loop gain in a op-amp will have no effect on the overall gain of a negative feedback system – *exactly what we require.*

We can now use this type of block diagram representation along with the more usual circuit diagrams to calculate the effects of feedback on the other amplifier parameters of input and output impedances, noise and distortion.

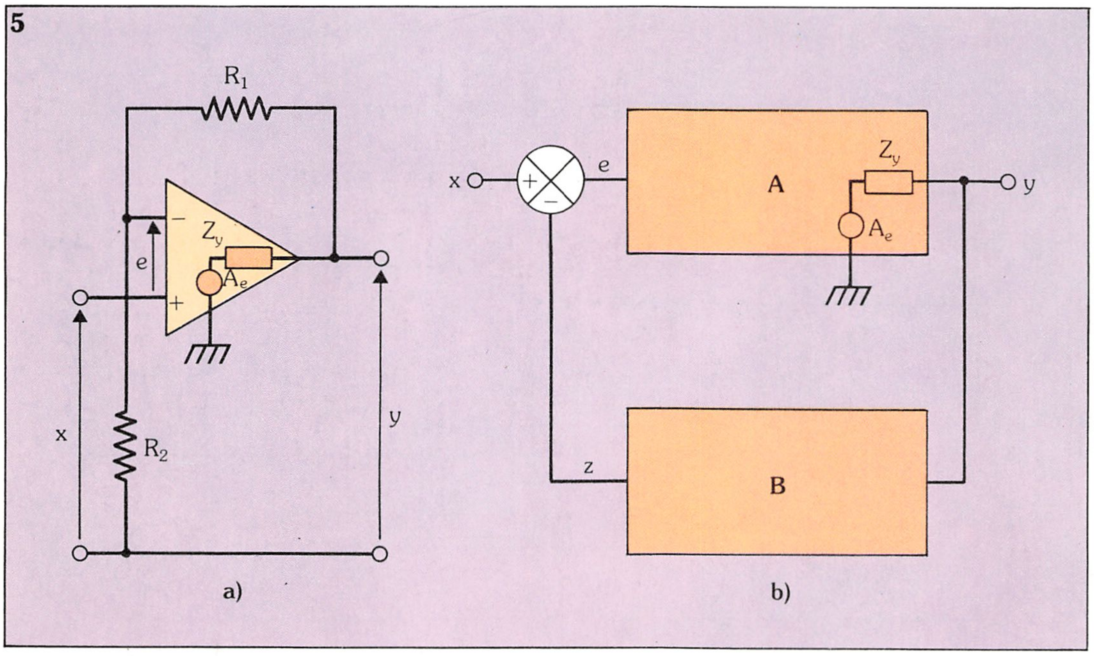
Input impedance of the non-inverting amplifier

In *figure 4a* a non-inverting amplifier circuit is shown with input voltage x, feedback voltage z, error voltage e, and output voltage y. An open loop input impedance, Z_e, is shown between the two op-amp input terminals. Also known is an input current, I_{in}. A block diagram representation of this circuit is shown in *figure 4b.*

The input current is the voltage across the open loop input impedance, divided by the impedance value:

$$I_{in} = \frac{e}{Z_e}$$

but: $e = x - z$



5. **(a)** A **non-inverting negative** feedback amplifier circuit; **(b)** its block diagram equivalent.

so:

$$I_n = \frac{x - z}{Z_e}$$

Now, the *closed loop* input impedance, Z_{in}, is equal to the input voltage, divided by the input current:

$$Z_{in} = \frac{x}{I_{in}}$$

$$= \frac{x Z_e}{x - z}$$

$$= \frac{Z_e}{1 - z/x}$$

We know that the gain:

$$\frac{y}{x} = \frac{A}{1 + AB}$$

and hence:

$$\frac{z}{x} = \frac{AB}{1 + AB}$$

So:

$$Z_{in} = \frac{Z_e}{1 - \dfrac{AB}{1 + AB}} = Z_e (1 + AB)$$

We can see that the input impedance of the op-amp, in a non-inverting feedback mode, is raised by the feedback factor.

Output impedance of the non-inverting amplifier

One way of calculating a circuit's output impedance is to divide the circuit's output voltage with no load (known as its **open circuit voltage**), by the output current when the output terminal is short circuited to ground (its **short circuit current**).

Figure 5a shows the circuit of a non-inverting negative feedback amplifier, whose output circuit is shown as a voltage generator and a series impedance. Output voltage of the generator is Ae, and the impedance is the open loop impedance, Z_y. A block diagram representation is shown in *figure 5b*. We can now see why the open circuit voltage, divided by the short circuit current, gives the value of the output impedance – because the open circuit voltage is the voltage across the impedance when the output is short circuited.

Now, the open circuit output voltage:

$$V_{oc} = Ae = \frac{Ax}{1 + AB}$$

The short circuit output current can be calculated because upon short circuit, $y = 0$. So the feedback signal, $z = 0$. Thus the error signal, $e = x$. So, the short circuit current:

$$I_{sc} = \frac{Ax}{Z_y}$$

The overall output impedance is thus:

$$Z_{out} = \frac{V_{oc}}{I_{sc}} = \frac{Ax Z_y}{(1 + AB) Ax}$$

$$= \frac{Z_y}{1 + AB}$$

The output impedance has been reduced by the feedback factor.

(*continued in part 25*)